
Krav på IA

Nästa Generation Modeller

Avancerad utbildning för handledare

Katalogprinciper

Verktyg

Informationspridning

Rapport K nr 1: IRDS

Rapport K nr 2: IRDS Modeller och modellnivåer

Rapport K nr 3: Koppling begreppsmodell - relationsmodell

Rapport K nr 4: IBM:s Repository Manager- en Introduktion

Rapport K nr 5: IBM:s Repository Manager: Datamodelleringsbegreppen

Rapport K nr 6: IBM:s Repository Manager: Begreppsmodellering i Information Model

Koppling begreppsmodell - relationsmodell

Stig Berild

Rapporten är skriven i och för TRIAD
delprojekt Katalogprinciper.

Spridningsförbehåll:

Denna rapport får endast
spridas och användas inom
de organisationer som deltar
som parter i TRIAD-projektet.

© TRIAD-parterna 1991

Innehåll

Diskussion kring koppling begreppsmodell - relationsmodell

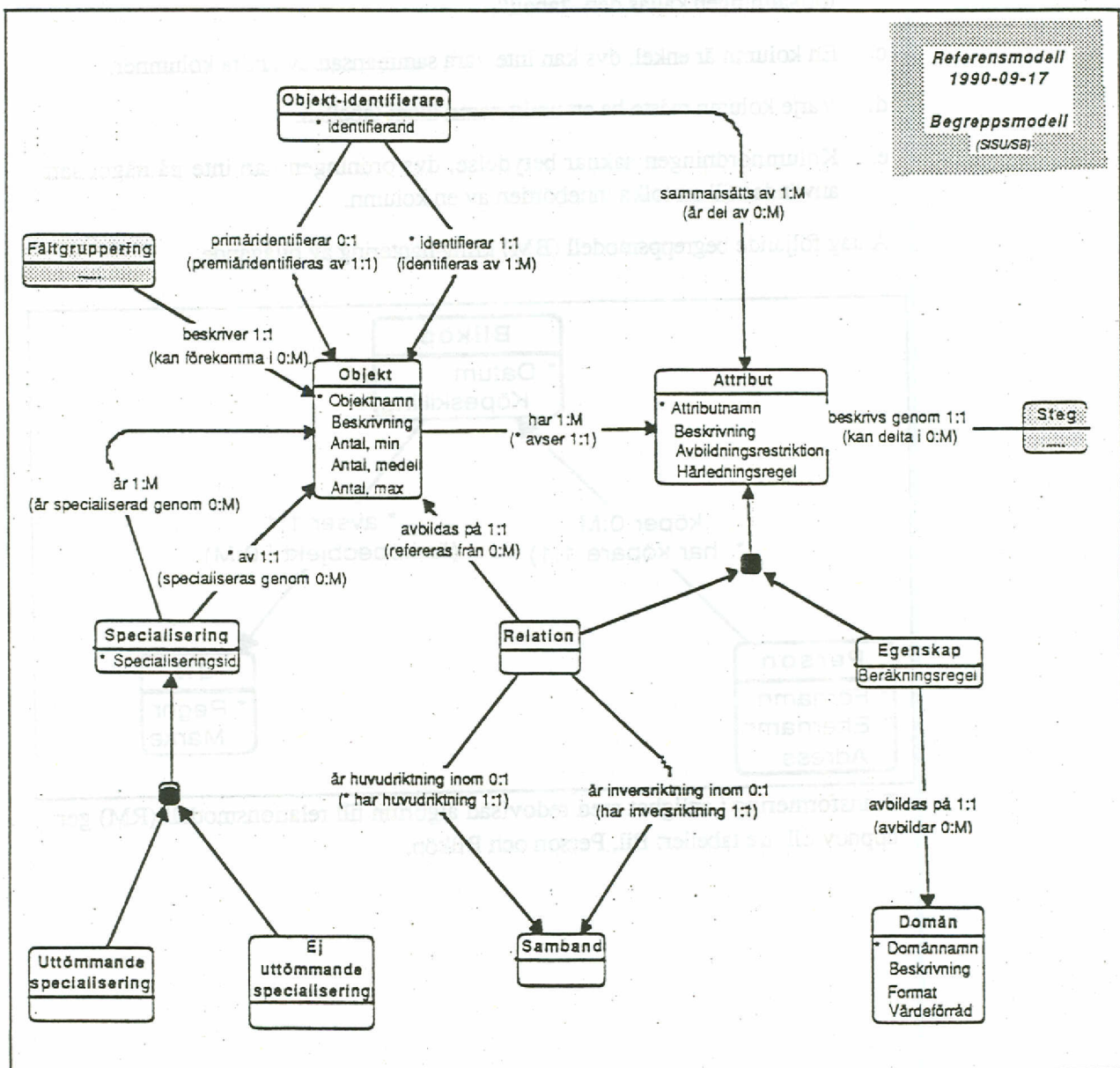
Prestandaanpassningar	15
Grova tumregler för transformation från begreppsmodell till relationsmodellen enligt 3:e normalform.....	17
Egenskapsfokusering	17
Sambandsfokusering.....	17
Specialiseringsfokusering	18

Diskussion kring koppling

begreppsmodell - relationsmodell

Denna rapport togs fram i anslutning till arbete inom Televerkets IA-projekt, specifikt "IA-Referensmodell", hösten 1990. Då rapporten bedömts ha relevans även för det pågående arbetet inom Triad, delprojekt Katalogprinciper, har Televerket givit sitt tillstånd till publicering inom Triad rapportserie.

Televerkets referensmodell strukturerar begreppsmodeller på följande sätt:

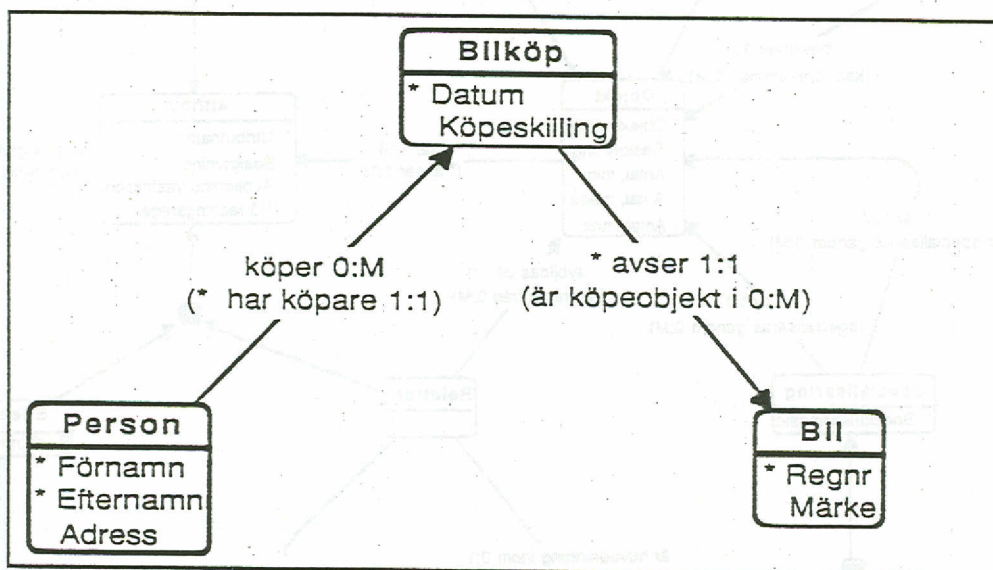


Hur bör en relationsmodell struktureras för att kunna upprätthålla intressanta kopplingar till den semantiskt rikare strukturen för begreppsmodeller, dvs den modell som den genererade relationsmodellen baseras på?

Vi utgår från de grova transformeringsregler mellan begreppsmodell och relationsmodell som redovisas i rapportens två sista sidor samt från följande förutsättningar:

- Med begreppsmodell menar vi här den del av en total begreppsmodell som avgränsats för att svara mot en logisk eller fysisk databas. En logisk databas kan bli horisontellt och eller vertikalt partitionerad av prestandaskäl eller för eventuell fysisk distribution över flera installationer. Både den logiska och den fysiska databasen i form av relationsmodeller antas alltså vara beskrivbara utifrån en avgränsning eller vy av den totala begreppsmodellen.
- En relation representeras av en tabell bestående av ett antal kolumner. I fortsättningen kallas den "tabell".
- En kolumn är enkel, dvs kan inte vara sammansatt av andra kolumner.
- Varje kolumn måste ha ett unikt namn inom tabellen.
- Kolumnordningen saknar betydelse, dvs ordningen kan inte på något sätt användas till att tolka innebörden av en kolumn.

Antag följande begreppsmodell (BM) kring hantering av bilägande.



Transformering i enlighet med redovisad algoritm till relationsmodell (RM) ger upphov till tre tabeller: Bil, Person och Bilköp.

Bil är enklast, med följande uppbyggnad:

Bil	Bil
Regnr (id)	Märke
Märke	Regnr (id)
(alt a)	(alt b)

Om namnet på tabellen sätts till namnet på motsvarande objekt i BM uppstår inga tolkningsproblem vad gäller tabellens innebörd. Det är heller inte svårt att se varifrån kolumnerna i (alt a) härör från i BM. Det blir inte svårare om vi byter plats på kolumnerna, som i (alt b). Deras namn refererar ändå klart till motsvarande begrepp i BM.

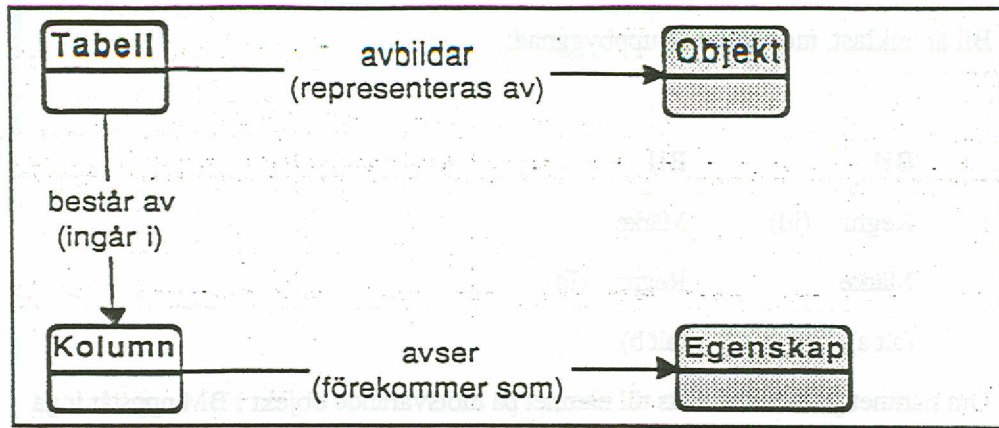
Om alltså både tabell och dess kolumner namnges i enlighet med BM har man dels en implicit koppling till BM, dels möjlighet att därifrån ta fram vilken eller vilka kolumner som bygger upp primärnyckeln (liksom eventuella alternativa nycklar).

Även person-objektet avbildas till en tabell som identifieras med två attribut respektive kolumner. Avbildningen ger sannolikt en tabell enligt (alt a).

Person	Person	Ägare
Förmamn (id)	Efternamn (id)	Tilltalsnamn (id)
Efternamn (id)	Adress	Bostad
Adress	Förmamn (id)	Familjenamn (id)
(alt a)	(alt b)	(alt c)

(Alt b) är likvärdigt eftersom namnen på kolumnerna är desamma som motsvarande egenskaper i begreppsmodellen.

Om man av någon anledning önskar använda andra namn för kolumner, som i (alt c), upphör den namnmässiga kopplingen med begreppsmodellen. Obs, att det visserligen finns anledning arbeta med så stor namnöverensstämmelse som möjligt mellan RM och BM men att det kan finnas fog för olika namn (vilket kommer att visa sig i kommande exempel). Avbildningsprinciperna och kopplingarna mellan RM och BM måste upprättas utan tanke på eventuell namnöverensstämmelse. Konsekvensen blir att både tabeller och kolumner i RM explicit måste peka ut sina respektive ursprung i BM. Se beskrivningsmodellen (metamodellen) nedan utgående från givna förutsättningar. Vilka kolumner som tillsammans utgör tabell-identifieringen, dvs primärnyckel, kan tills vidare härledas via BMs beskrivning av vilka egenskaper som ingår som identifierare.

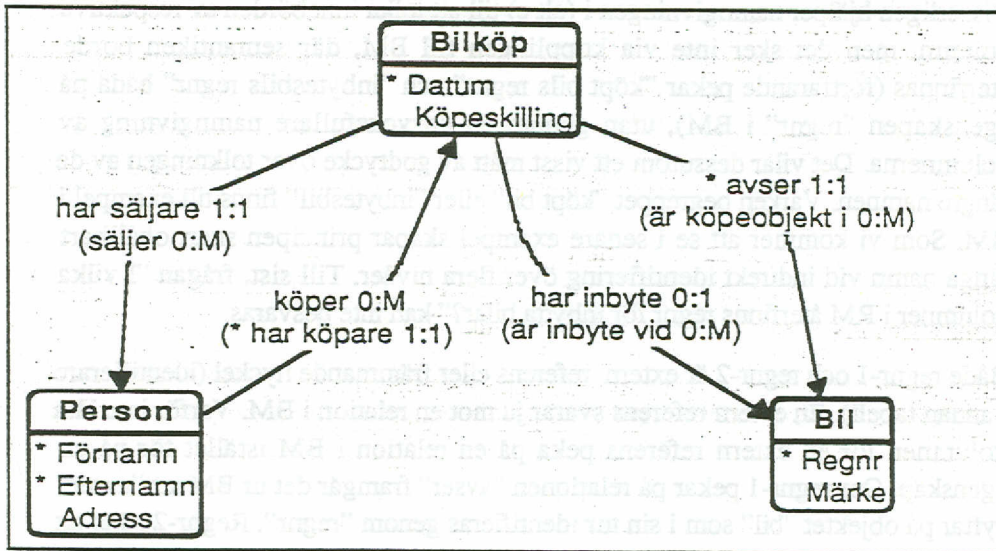


Den tredje tabellen, bilköp, får en uppbyggnad enligt (alt a), om man inte ändrar i avbildningsalgoritmens förslag. Eftersom många omständigheter (namngivningsregler på RM-nivå, prestanda ...) kan motivera förändringar av förslaget måste även en tabell enligt ex.vis (alt b) klaras av.

Bilköp		Ägarskap	
Förmamn	(id)	Ikraftträdande	(id)
Efternamn	(id)	E-namn	(id)
Datum	(id)	Registreringsnummer	(id)
Regnr	(id)	Belopp	
Köpeskilling		F-namn	(id)
(alt a)		(alt b)	

Det nya i denna tabell är att delar av identifieringen hämtas över relationer i BM från andra objekts identifikationer. Med den givna metamodellen kan man fortfarande med viss möda härleda deras motsvarigheter i BM. Kolumnen "f-namn" pekar på person-egenskapen "förmamn". Person är i sin tur relaterad via "köper" med "bilköp". Alltså är "f-namn" förmamnet på den person som är köparen vid bilköpet.

Utvidgar vi exemplet något fungerar inte längre härledningsstrategin. Ta följande utvidgning av bilköpet:



Alternativa motsvarande tabell-utformningar:

Bilköp	Bilköp	Bilköp
Förmamn (id)	Förmamn-1 (id)	Köparens förmamn (id)
Efternamn (id)	Efternamn-1 (id)	Köparens efternamn (id)
Datum (id)	Datum (id)	Datum (id)
Regnr (id)	Regnr-1 (id)	Köpt bils regnr (id)
Köpeskilling	Köpeskilling	Köpeskilling
Förmamn	Förmamn-2	Säljarens förmamn
Efternamn	Efternamn-2	Säljarens efternamn
Regnr	Regnr-2	Inbytesbils regnr
(alt a)	(alt b)	(alt c)

(Alt a) är en direktöversättning av egenskapsnamnen, men strider mot förutsättning d ovan om unika kolumnnamn. (Alt b) har tillförts suffix för att klara av namnkravet och är en korrekt tabell. Men nu tillstöter problem. Både regnr-1 och regnr-2 pekar ut samma egenskap "regnr" i BM. Inget säger att

regnr-1 hör till just den köpta bilen och inte till inbytesbilen. Det kunde lika gärna vara regnr-2, som då skulle noteras som delidentifierare istället för regnr-1.

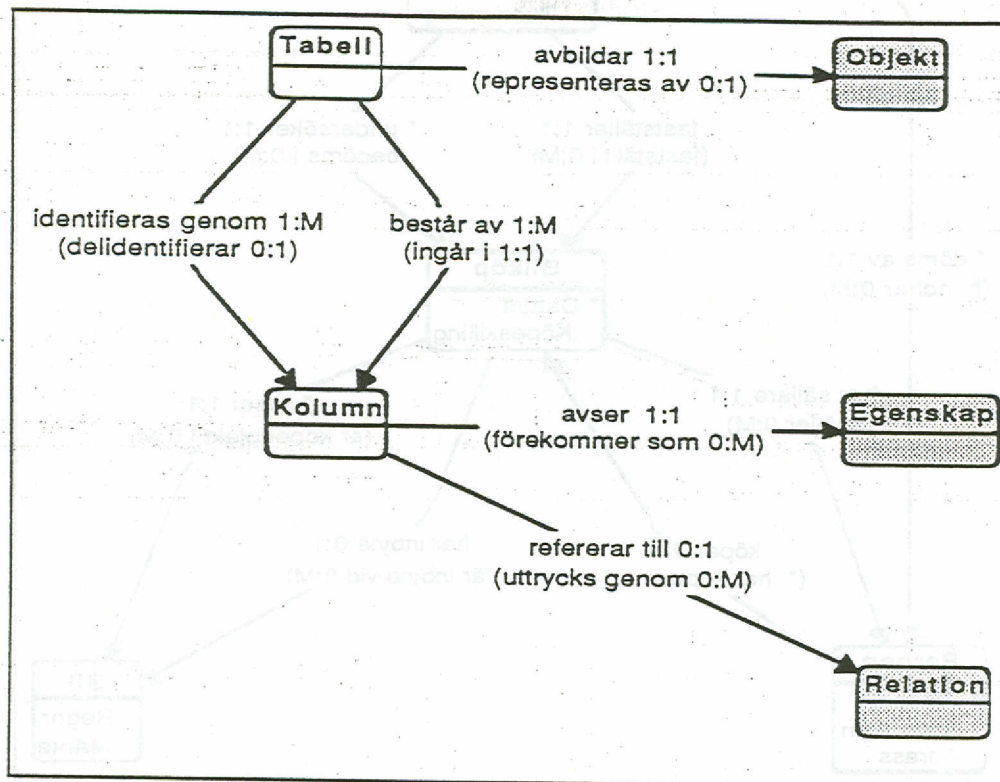
Visserligen hjälper namngivningen i (alt c) till att tolka innebörden av respektive kolumn, men det sker inte via kopplingen till BM, där semantiken borde återfinnas (fortfarande pekar "köpt bils regnr" och "inbytesbils regnr" båda på egenskapen "regnr" i BM), utan genom en uttrycksfullare namngivning av kolumnerna. Det vilar dessutom ett visst mått av godtycke över tolkningen av de längre namnen. Varken begreppet "köpt bil" eller "inbytesbil" finns till exempel i BM. Som vi kommer att se i senare exempel skapar principen snart ohållbart långa namn vid indirekt identifiering över flera nivåer. Till sist, frågan "I vilka kolumner i RM återfinns regnr för inbytta bilar?" kan inte besvaras.

Både regnr-1 och regnr-2 är extern referens eller främmande nyckel (identifierare i annan tabell). En extern referens svarar ju mot en relation i BM. Varför inte låta kolumnen för en extern referens peka på en relation i BM istället för på en egenskap. Om regnr-1 pekar på relationen "avser" framgår det ur BM att "avser" syftar på objektet "bil" som i sin tur identifieras genom "regnr". Regnr-2 sätts att peka på "har inbyte" och entydighet har uppnåtts. Vilka kolumner som deltar i identifieringen av en tabell kan nu också entydigt härledas ur BM. Som komplement kan metamodellen tillföras sambandet "identifieras genom", som direkt pekar ut dessa kolumner.

Är samma lösning överförbar på bilköpets köpare och säljare? Ja, till viss del. Även här gäller det att peka ut rätt relation i BM istället för egenskaper. Skillnaden ligger i att den externa nyckeln består av två kolumner i RM. Låter vi förmamn-1 och eftermamn-1 båda peka på relationen "har köpare" är det entydigt att dessa två tillsammans identifierar den köpande personen. Risker för sammanblandning med förmamn-2 eller eftermamn-2 existerar inte. Vad som dock inte framgår är att förmamn-1 står för personens egenskap "förmamn" och eftermamn-1 för "eftermamn". Observera, återigen, att vi i resonemanget måste bortse från att namnen på kolumn och egenskap är snarlika. Vi måste dels tala om vilken egenskap kolumnen motsvarar, dels vilken relation de båda kolumnerna tillsammans svarar mot.

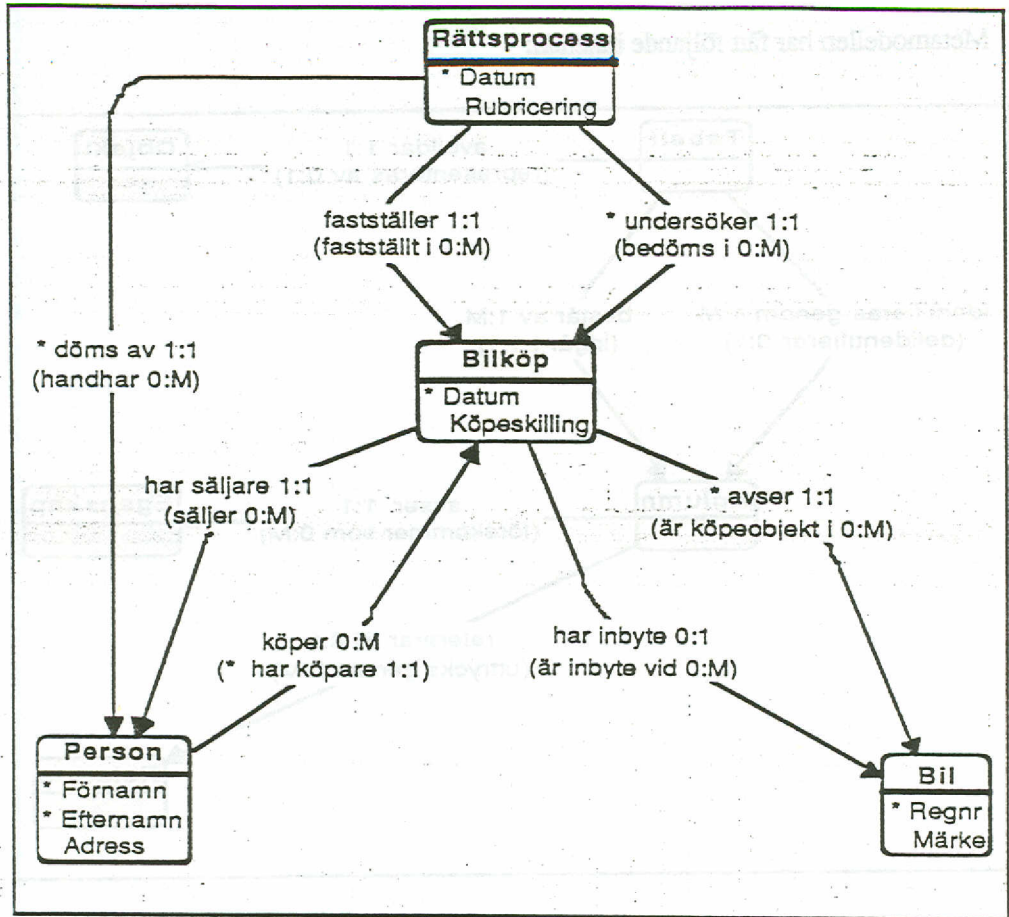
Det blir ingen synlig ändring i metamodellen, bara en vetskap om att flera kolumner ur samma tabell nu kan referera till samma relation. Därtill kommer att en och samma kolumn ibland pekar på både en egenskap och en relation.

Metamodellen har fått följande innehåll:



Det som nu utretts är referens mellan egenskaper och kolumner samt mellan relationer på en nivå och deras kolumnmotsvarigheter. Är metamodellen därmed även generellt heltäckande?

Vi testar genom att komplettera begreppsmodellen enligt nedan. Exemplet kan måhända upplevas som konstruerat men saknar inte realism. Denna typ av situation liksom ännu mer komplexa dito måste kunna hanteras i en metamodell.



Objektet "rättsprocess" ger upphov till en av följande tre alternativa tabeller:

Rättsprocess	Rättsprocess	Rättsprocess
Förnamn-1 (id)	Köparens förnamn (id)	Undersökt ägarskaps köparens förnamn (id)
Efternamn-1 (id)	Köparens efternamn (id)	Undersökt ägarskaps köparens efternamn (id)
Datum-1 (id)	Bilköpsdatum (id)	Undersökt ägarskaps datum (id)
Regnr-1 (id)	Köpt bils regnr (id)	Undersökt ägarskaps bils regnr (id)
Förnamn-2 (id)	Domarens förnamn (id)	Domarens förnamn (id)
Efternamn-2 (id)	Domarens efternamn (id)	Domarens efternamn (id)
Datum (id)	Rättsprocessdatum (id)	Rättsprocessdatum (id)
Förnamn-3	Köparens förnamn-2	Fastställt ägarskaps köparens förnamn
Efternamn-3	Köparens efternamn-2	Fastställt ägarskaps köparens efternamn
Datum-3	Bilköpsdatum-2	Fastställt ägarskaps datum
Regnr-3	Köpt bils regnr-2	Fastställt ägarskaps bils regnr
Rubricering	Rubricering	Rubricering

(alt a)

(alt b)

(alt c)

(Alt a) ger den torftigaste förståelsen. Vi har exempelvis så mycket som tre stycken förmamn som vi inte vet exakt vad de står för. Även om vi antar att (id)-angivelserna hamnat rätt finns fortfarande två förmamn kvar inom identifieringen, vars roller inte framgår. (Alt b) svarar ungefär mot den semantik som är uttryckbar genom metamodellen.

I det tidigare exemplet lät vi för- och efternamnen, förutom på egenskaper i BM, även peka på den relation som motiverat deras respektive medverkan i relationen. En kombination för- och efternamn pekade på "har köpare", den andra på "har säljare". Nu räcker inte den lösningen längre. Vi har två kolumner som båda pekar på egenskapen "förmamn" och på relationen "har köpare". Samma sak gäller för efternamn. Tillsammans pekar alltså fyra kolumner på "har köpare".

Orsaken finns att hämta i de båda roller "bilköp" befinner sig under "rättsprocess", nämligen både som undersökt och som fastställt bilköp. I båda fallen del-identifieras bilköpet med identifieraren för köparen. Det måste kunna gå att uttrycka vilket par av för- och efternamn som hänvisar till köparen i det undersökta bilköpet och vilket par som gäller för köparen i det fastställda bilköpet. (Alt c) ovan visar en semantik i kolumn-benämningen som i praktiken måste kunna framgå genom referenserna till BM. Vi har redan tidigare konstaterat att det inte går att förlita sig på namngivningen av kolumner. Det kan t.ex. säkert finnas situationer som är betydligt mer komplexa än denna.

Entydighet kan bara uppnås om varje relation kopplas till den kombination kolumner som svarar mot den främmande nyckel relationen uttrycks genom i en tabell. Vi måste exempelvis skapa den grupp kolumner som utgör den främmande nyckeln för bilköp i rollen av undersökt bilköp. Tjuvtittar vi på alternativet c men väljer benämningarna enligt alternativ a (för att ta det mest intetsägande alternativet) får vi gruppen:

förmamn-1

efternamn-1

regnr-1

datum-1

men det skulle lika gärna kunnat vara (men då med annan placering av id-noteringar):

förmamn-3

efternamn-2

regnr-1

datum

Motsvarande gruppering måste göras för fastställt bilköp. Vi kan låta fyra kolumner peka på relationen "fastställer" och fyra andra på relationen "undersöker" enligt tidigare strategi. Dock finns inom respektive gruppering två andra främmande nycklar, nämligen den som pekar ut person i rollen som köpare och bil i rollen av köpt objekt. Ska ex.vis kolumnen förmamn-1 peka både på "undersöker" och "har köpare"? Här räcker det med "undersöker" eftersom det i detta exempel inte finns någon risk att inom den större grupperingen missförstå betydelsen av förmamn. Vi har bara ett förmamn inom gruppen. Det kan entydigt härledas till köparen.

Därmed inte sagt att det inte kan finnas situationer där flera förmamn uppträder inom gruppen liksom del-grupperingar med samma egenskapsreferenser. Exempel på detta är grupperingen av de kolumner som utgör identifieraren för en rättsprocess. Denna kan ju tänkas ingå som delidentifierare i något annat sammanhang. Semantiken måste kunna framgå entydigt även i de mest komplicerade fallen. Nyckeln ser ut som följer:

Identifierare/primärnyckel

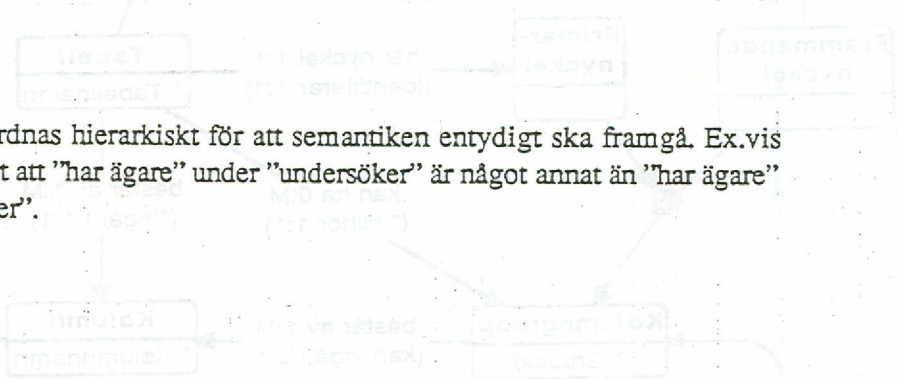
för rättsprocess	undersöker	har ägare	avser döms av
förmamn-1	förmamn-1	förmamn-1	
eftermamn-1	eftermamn-1	eftermamn-1	
regnr-1	regnr-1		regnr-1
datum-1	datum-1		
förmamn-2			förmamn-2
eftermamn-2			eftermamn-2
datum			

Där nyckeln figurerar som en främmande nyckel räcker det inte att låta dess kolumner peka på den relation som givit upphov till den främmande nyckeln. Vi måste inom nyckeln skilja ut ex.vis rollen för förmamn-1 från rollen för förmamn-2. Subgruppering av referenser till "undersöker" respektive "döms av" blir nödvändig. Resultatet blir att förmamn-1 ingår i primärnyckeln för rättsprocess och i sub-gruppen som pekar på "undersöker". Förmamn-2 ingår i samma primärnyckel samt i sub-gruppen som pekar på "döms av". För att slippa arbeta med härledning (där detta går) väljer vi att konsekvent skapa en kolumngrupp för varje främmande nyckel. Kolumngruppen för "har ägare" under nyckeln skulle ex.vis för närvarande inte behövas, men vad säger att denna situation för alltid består?

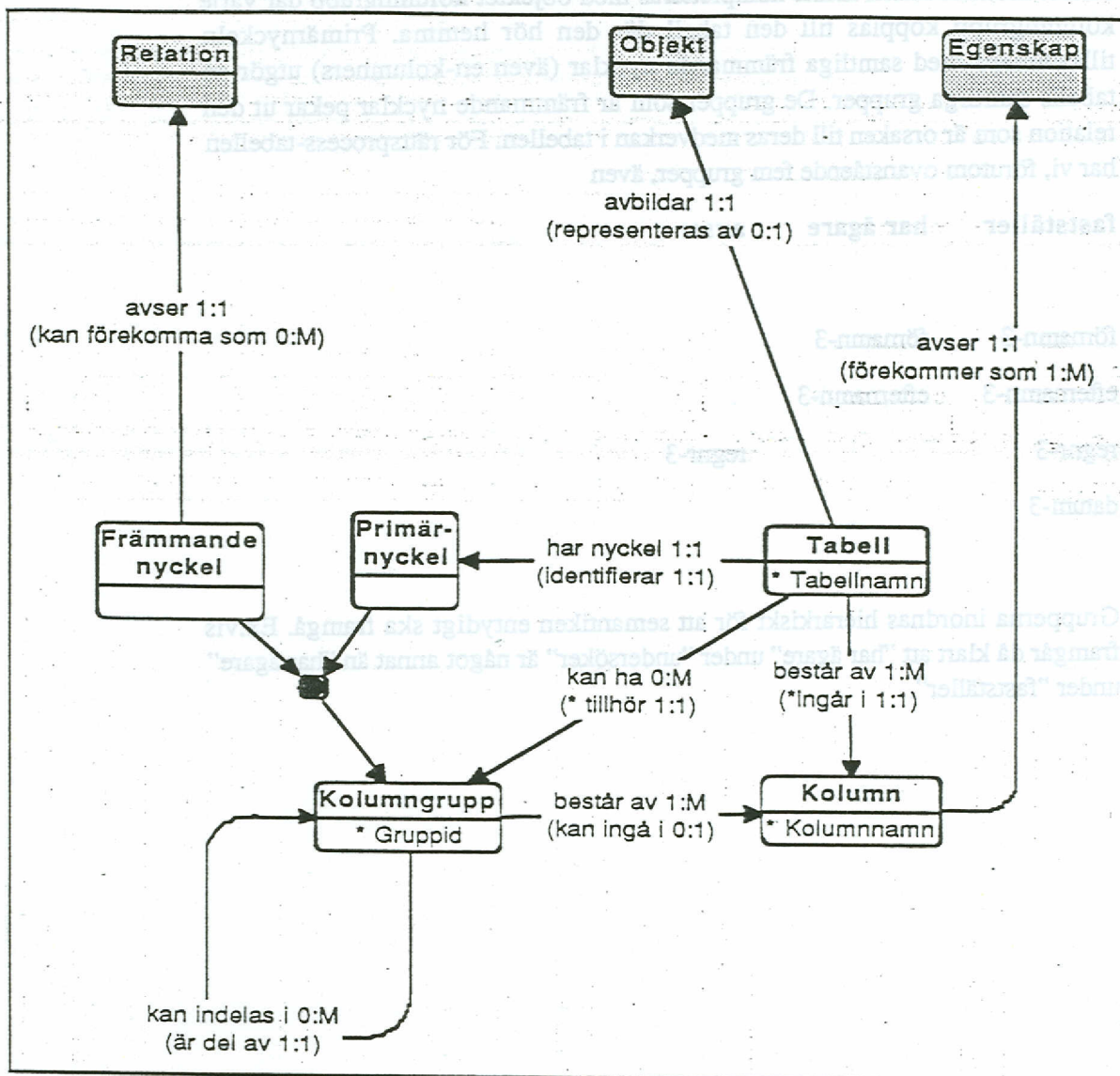
M.a.o. metamodellen måste kompletteras med objektet kolumngrupp där varje kolumngrupp kopplas till den tabell där den hör hemma. Primäryckeln tillsammans med samtliga främmande nycklar (även en-kolumners) utgör en tabells samtliga grupper. De grupper som är främmande nycklar pekar ut den relation som är orsaken till deras medverkan i tabellen. För rättsprocess-tabellen har vi, förutom ovanstående fem grupper, även

fastställer har ägare avser
 förmamn-3 förmamn-3
 efternamn-3 efternamn-3
 regnr-3 regnr-3
 datum-3

Grupperna inordnas hierarkiskt för att semantiken entydigt ska framgå. Ex.vis framgår då klart att "har ägare" under "undersöker" är något annat än "har ägare" under "fastställer".

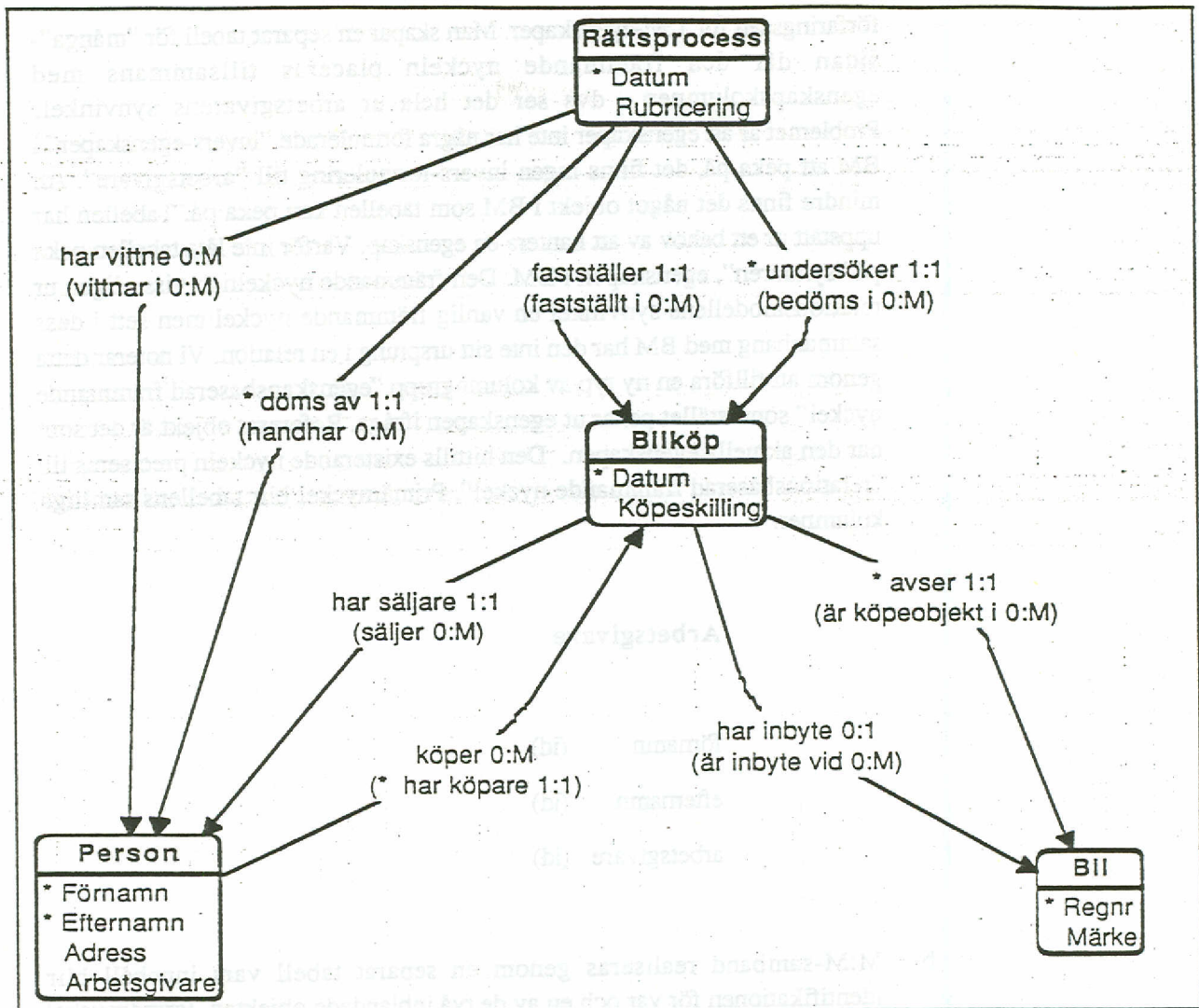


Metamodellen presenteras nedan i sitt kompletterade skick.



Till sist belyser vi hur de tidigare presenterade transformeringsreglerna påverkar kopplingen mellan BM och RM. Inkluderat i den existerande rättsprocessmodellen är 1:M-samband (ex.vis över relationen "köper"), M:1-samband (ex.vis över relationen "undersöker") och envärda egenskaper. 1:1-samband kräver bedömning kring var i de två tabellerna den främmande nyckeln ska placeras men hanteras annars i princip enligt M:1-samband. Kvår att behandla är flervärda egenskaper och M:M-samband.

För att beskriva dessa två varianter kompletteras rättsprocessmodellen enligt följande:



Tillfogat är följande:

- a. Egenskapen "arbetsgivare" till objektet person. En person kan ha flera arbetsgivare.
 - b. De personer som inkallas som vittnen under processen. En process kan ha flera vittnen. En person kan vara vitne vid flera processer.
- a. Väljer man att stanna på 1:a normalformen blir det inga andra ändringar än att tabellen för person kompletteras med en kolumn för arbetsgivare. Kolumnen pekar ut egenskapen "arbetsgivare" i BM.

Om målet är tabeller enligt 3:e normalformen behöver denna 1:M-egenskap hanteras speciellt. Vid 1:M-samband placerar vi enkelt den främmande nyckeln på "många"-sidan, dvs med M:1-aspekten, samt sätter den att peka på sambandets M:1-relation i BM. I grunden skulle man önska samma

förfaringssätt för 1:M-egenskaper. Man skapar en separat tabell för "många"-sidan där den främmande nyckeln placeras tillsammans med egenskapskolumnen, dvs ser det hela ur arbetsgivarens synvinkel. Problemet är att egenskaper inte har några formulerade "invers-egenskaper" i BM att peka på, det finns ingen invers-formulering till "arbetsgivare". Än mindre finns det något objekt i BM som tabellen kan peka på. Tabellen har uppstått ur ett behov av att hantera en egenskap. Varför inte låta tabellen peka på "syndaren", egenskapen i BM. Den främmande nyckeln är visserligen ur relationsmodellens synvinkel en vanlig främmande nyckel men sett i dess sammanhang med BM har den inte sitt ursprung i en relation. Vi noterar detta genom att tillföra en ny typ av kolumngrupp "egenskapsbaserad främmande nyckel" som istället pekar ut egenskapen ifråga. Refererat objekt är det som har den aktuella egenskapen. Den hittills existerande nyckeln preciseras till "relationsbaserad främmande nyckel". Primärnyckel blir tabellens samtliga kolumner.

Arbetsgivare

förmamn (id)

eftermamn (id)

arbetsgivare (id)

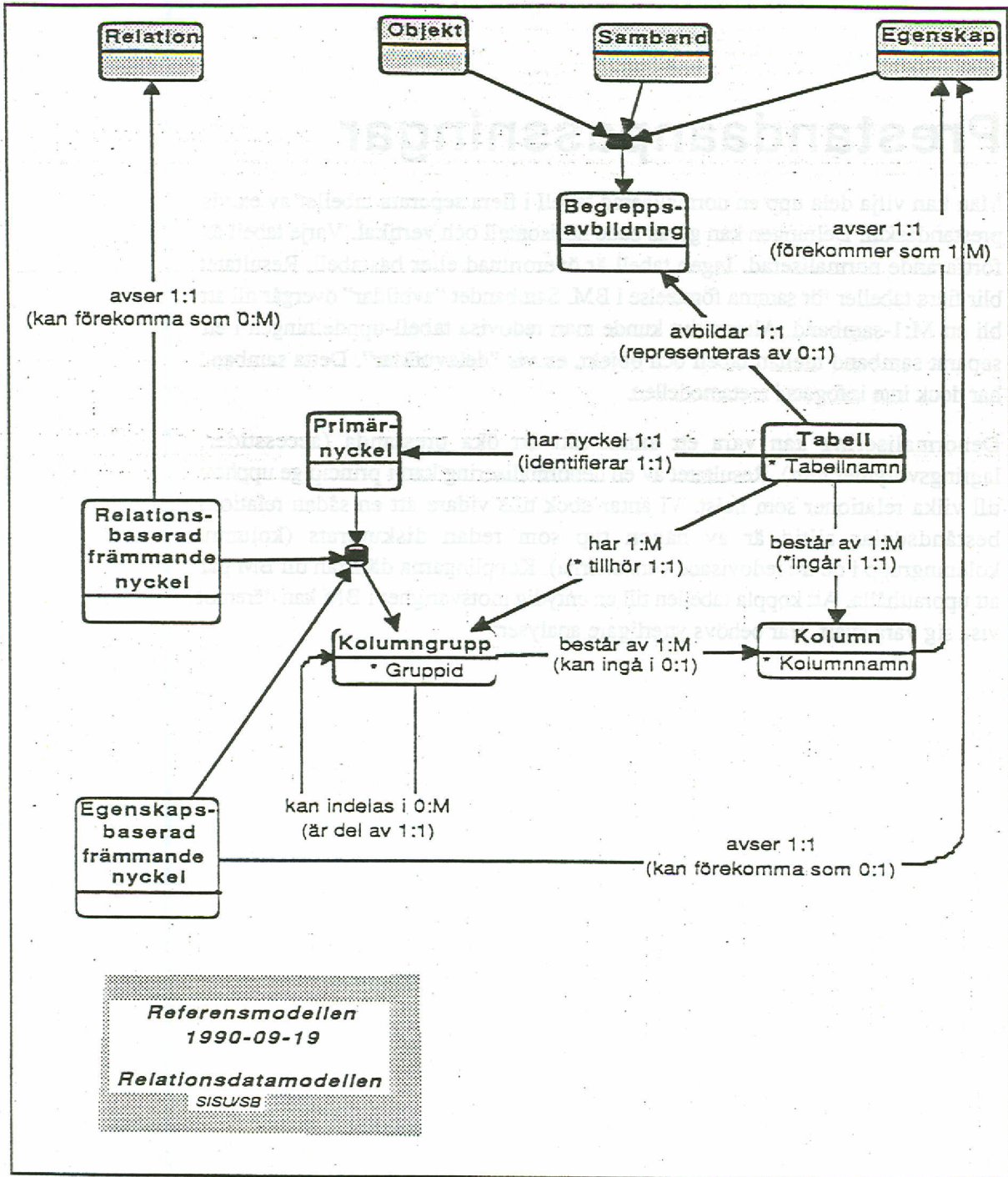
- b. M:M-samband realiseras genom en separat tabell vars innehåll blir identifikationen för var och en av de två inblandade objekten. Primärnyckel är summan av dessa två kolumngrupper. Var och en av de två kolumngrupperna är en främmande nyckel som sätts att peka på den relation som står för ett M:1-förhållande till objektet. I exemplet pekar den främmande nyckeln för "rättsprocess" på "vittnar i" medan nyckeln för "person" pekar på "har vittne". Det är lättare att se situationen om man upplever tabellen som en representation av "person som vittne i viss rättsprocess". I dennes specifika perspektiv finns ju bara en person, som återfinns över "har vittne" och en rättsprocess, som återfinns över "vittnar i".

Den enda komplikationen uppstår vid relateringen av tabellen till dess motsvarighet i BM. Ursprunget är inte ett objekt utan ett samband.

Prestandaanpassningar

Man kan vilja dela upp en normaliserad tabell i flera separata tabeller av ex.vis prestandaskäl. Delningen kan göras både horisontell och vertikal. Varje tabell är fortfarande normaliserad. Ingen tabell är överordnad eller bastabell. Resultatet blir flera tabeller för samma företeelse i BM. Sambandet "avbildar" övergår till att bli ett M:1-samband. Alternativt kunde man redovisa tabell-uppdelningen i ett separat samband mellan tabell och objekt, ex.vis "delavbildar". Detta samband har dock inte infogats i metamodellen.

Denormalisering kan vara ett annat sätt att öka prestanda (accesstider, lagringsvolym mm). Resultatet av en denormalisering kan i princip ge upphov till vilka relationer som helst. Vi antar dock tills vidare att en sådan relations beståndsdelar alltid är av någon typ som redan diskuterats (kolumn, kolumngrupp i de tre redovisade versionerna). Kopplingarna därifrån till BM går att upprätthålla. Att koppla tabellen till en entydig motsvarighet i BM kan däremot visa sig vara svårt. Här behövs ytterligare analyser.



Grova tumregler för transformation från begreppsmodell till relationsmodellen enligt 3:e normalform.

Denna kortfattade utredning har gjorts med syftet att utröna vilka kopplingar mellan begreppsmodell och relationsmodell som kan bli aktuella att avbilda i referensmodellen.

Egenskapsfokusering

- a. För varje objekt som har envärda egenskaper (E), skapa tabell med objektets primärnyckel (OPN) samt de envärda egenskaperna. (OPN, E1,...,En)

Ev. optimering: Undvik många NULLs genom att bryta ut NULL-rika egenskaper (eller egenskapsgrupper) till egna tabeller. (OPN, E1,...,Ex),..., (OPN, Ey),..., (OPN, Ez,...En)

- b. För varje flervärd egenskap, skapa tabell för varje sådan egenskap innehållande objektets primärnyckel och egenskapen. (OPN,E)

Sambandsfokusering

- c. För varje 1:1-samband av typen

0,1:0,1 placera objektnyckeln för objektet med störst kardinalitet i tabellen för objektet med minst kardinalitet som främmande nyckel.

$O1 \text{ ---} 0,1 \text{ ---} s \text{ ---} 0,1 \text{ ---} O2$ blir $O1: (OPN1, OPN2, \dots)$
 $O2: (OPN2, \dots)$

100 100000

Ev. optimering: Om många NULLS i referensen, placera sambandet i separat tabell. (OPN1,...), (OPN2,...), (OPN1, OPN2)

0,1:1,1 placera främmande nyckeln i tabellen för objektet med 1:1-koppling för att undvika NULLS.

$O1 \text{ ---} 0,1 \text{ ---} s \text{ ---} 1,1 \text{ ---} O2$ blir $O1: (OPN1, \dots)$
 $O2: (OPN2, OPN1, \dots)$

1,1:1,1 välj endera objekt-tabellen.

$O1: (OPN1, OPN2, \dots)$ och $O2: (OPN2, \dots)$

eller

$O1: (OPN1, \dots)$ och $O2: (OPN2, OPN1, \dots)$

- d. För varje 1:M- eller M:1-samband, placera nyckeln på tabellen för objektet på M-sidan.

$O1 \text{ --- } x,1 \text{ --- } s \text{ --- } x,M \text{ --- } O2$ blir $O1: (OPN1, \dots)$
 $O2: (OPN2, OPN1, \dots)$

($x=0$ eller 1)

Ev. optimering: Om många NULLS i främmande nyckeln, placera sambandet i separat tabell.

$O1 \text{ --- } 0,1 \text{ --- } s \text{ --- } x,M \text{ --- } O2$ blir $(OPN1, \dots), (OPN2, \dots)$
 $(OPN1, OPN2)$

($x=0$ eller 1)

- e. För varje M:M-samband skapas en separat tabell med båda objektens nycklar.

$O1 \text{ --- } x,M \text{ --- } s \text{ --- } x,M \text{ --- } O2$ blir $(OPN1, OPN2)$

Specialiseringsfokusering

- f. För super/sub-samband om

super har id	se i princip 1:1-samband ovan där åtminstone 0,1(super):1,1(sub) gäller. Har sub eget id används detta annars ärvs supers id.
inget superid	implicerar olika id på subnivå. Som ovan, men först måste artificiell id skapas för super.

Optimeringar därutöver

är beroende av

- vald teknisk lösning
- tänkta framtida anpassningar och utvidgningar av modell
- systemsamordning

mm

och behandlas därför inte vidare i detta forum.

TRIAD utvecklar IA

Televerket har just tagit första steget in i sin nya IA-organisation och Posten håller på att bygga upp sin nya DA-organisation. Båda organisationerna har sett nytan att inför 90-talet gå vidare tillsammans i TRIAD-projektet som drivs tillsammans med SISU. Statskontoret deltar också i projektet för att på sikt kunna föra ut nya synsätt och hjälpmedel inom den civila statliga sektorn.

Ericsson Data Services deltar med tyngdpunkten i den del som handlar om att utveckla kompetenta modelleringsledare, delprojektet "Avancerad utbildning för modelleringsledare".

Modelleringsmetoder är centrala i bedrivandet av verksamheten inom informationsadministrationen. Därför arbetar ett delprojekt med utvecklandet av "nästa generation modelleringsmetod" som skall sättas i händerna på informationsadministratören. Siktet är att fördjupa och bredda dagens modelleringsmetoder och där hämta in kunskap från pågående forskning och utveckling internationellt. (faktaruta om IAS91).

Som stöd för informationsadministrationen behövs verktyg. Inom TRIAD arbetar man där inom två områden, kataloger och verktyg.

Delprojektet kataloger arbetar dels med att utforma den informationsmodell som måste kunna täckas av en katalog, dels med att granska och följa utvecklingen av produkter inom området t ex IBM:s "Repository" och Digital's "CDD". Dessutom följer man standardiseringen internationellt kring IRDS. För parterna i projektet liksom för andra organisationer är detta ett tungt område både vad gäller kommande investeringar ekonomiskt och vad gäller kompetenta resurser för en kommande övergång till "repository-världen". - Det inledande skedet syftar till att bygga upp en kunskapsplattform, som sedan kommer att kunna utnyttjas för kravställande och planering och genomförande av övergång från dagens kataloghantering till morgondagens.

Den andra verktygshanterande delen inom TRIAD-projektet, delprojektet "verktyg för informationsadministration", syftar till att ta fram verktyg för uttag och dokumentering av modeller. Betoningen ligger på människa datorgränssnitt och i första skedet görs utveckling av HYBRIS-gränssnittet med prototyper för Posten och för Televerket.

För att hålla ett helhetsperspektiv på projektets delar och för att ha inpassningen av funktionen Informationsadministration i organisationens övriga verksamhet arbetar delprojektet "Krav på IA". I delprojektet arbetar man dels med att kartlägga dagens krav på dataadministration och projicera till morgondagens krav på IA. Dessutom skall man skapa en bild av IA-verksamhetens innehåll och organisation. Från detta i sin tur ställer man krav

på övriga delprojekt. Vilka krav skall ställas på kompetens, metoder, hjälpmedel typ kataloger och gränssnitt?

TRIAD projektet är stort

Budgeten för TRIAD-projektet löper på 10 MSEK per år under en treårsperiod som startar vid kalenderåret 1991 års början och som alltså beräknas avslutad vid utgången av 1993.

TRIAD-projektet är ett tillämpningsprojekt

Det innebär att parterna, Televerket, Posten, Statskontoret, EDS och SISU går in med såväl persontidsansatningar som ekonomiska och att STU, Styrelsen för Teknisk Utveckling, bidrar med ett ekonomiskt tillskott som svarar mot ungefär 40 % av den insatta persontiden.

Öppet för fler deltagare

Parterna i TRIAD-projektet vill gärna öka tempot och bredda perspektivet och vill därför gärna ha fler parter in i projektet. Dessa parter får då enligt SISU:s tårtprincip "betala för en tårbit, men ät hela tårtan", tillgång till projektets resultat med en insats som ger stor "price performance".

Nya deltagare kan gå in i hela projektet eller i det eller de delprojekt som verkar intressantast. En förutsättning är att man framförallt är beredd att satsa kompetent personal. För de flesta intressenter bord detta vara ett utmärkt sätt att driva personalutveckling för personer t ex inom DA-området, samtidigt som man bygger upp beredskapen inför 90-talets IA-verksamhet.

Kompetensutveckling viktigt resultat

En viktig effekt för parterna av deras medverkan i TRIAD är kompetensutveckling. Man satsar på att ta in personer som så småningom eller redan idag arbetar med DA och IA för att ge dem en djup och "frontlinje"-mässig kompetens. Detta skall utnyttjas när man successivt för in resultaten i den egna organisationen. Projektdeltagarna har alltså en viktig roll som kunskapsförmedlare i den egna organisationen. Dessutom ger projektarbetet deltagarna tillfälle till en egen utveckling inom det professionella området som är unik.

Informationsspridning

Det sjätte delprojektet "Informationsspridning" har till uppgift att sörja för att i första hand parterna men också SISU:s övriga intressenter successivt kan följa och tillgodogöra sig resultat från TRIADprojektet. Seminarier, rapporter och referensgruppsverksamhet är led i den verksamheten.

Krav på IA

Nästa Generation Modellering

Avancerad utbildning för handledare

Katalogprinciper

Verktyg

Informationspridning

Rapport K nr 1: IRDS

Rapport K nr 2: IRDS Modeller och modellnivåer

Rapport K nr 3: Koppling begreppsmodell - relationsmodell

Rapport K nr 4: IBM:s Repository Manager- en Introduktion

Rapport K nr 5: IBM:s Repository Manager: Datamodelleringsbegreppen

Rapport K nr 6: IBM:s Repository Manager: Begreppsmodellering i Information Model

Koppling begreppsmodell - relationsmodell

Stig Berild

Rapporten är skriven i och för TRIAD
delprojekt Katalogprinciper.

Spridningsförbehåll:

Denna rapport får endast
spridas och användas inom
de organisationer som deltar
som parter i TRIAD-projektet.

© TRIAD-parterna 1991

Innehåll

Diskussion kring koppling begreppsmodell - relationsmodell

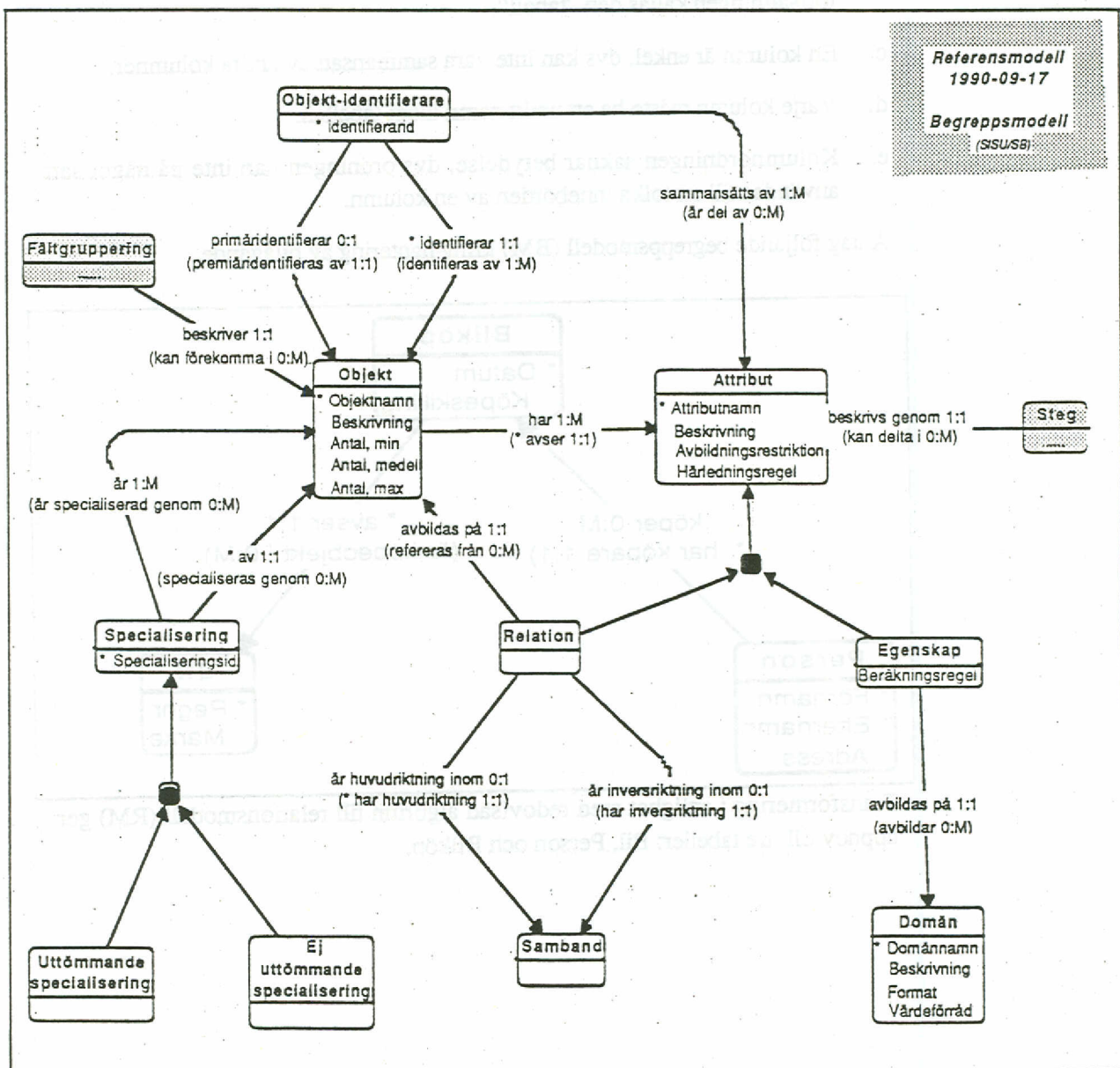
Prestandaanpassningar	15
Grova tumregler för transformation från begreppsmodell till relationsmodellen enligt 3:e normalform.....	17
Egenskapsfokusering	17
Sambandsfokusering.....	17
Specialiseringsfokusering	18

Diskussion kring koppling

begreppsmodell - relationsmodell

Denna rapport togs fram i anslutning till arbete inom Televerkets IA-projekt, specifikt "IA-Referensmodell", hösten 1990. Då rapporten bedömts ha relevans även för det pågående arbetet inom Triad, delprojekt Katalogprinciper, har Televerket givit sitt tillstånd till publicering inom Triad rapportserie.

Televerkets referensmodell strukturerar begreppsmodeller på följande sätt:

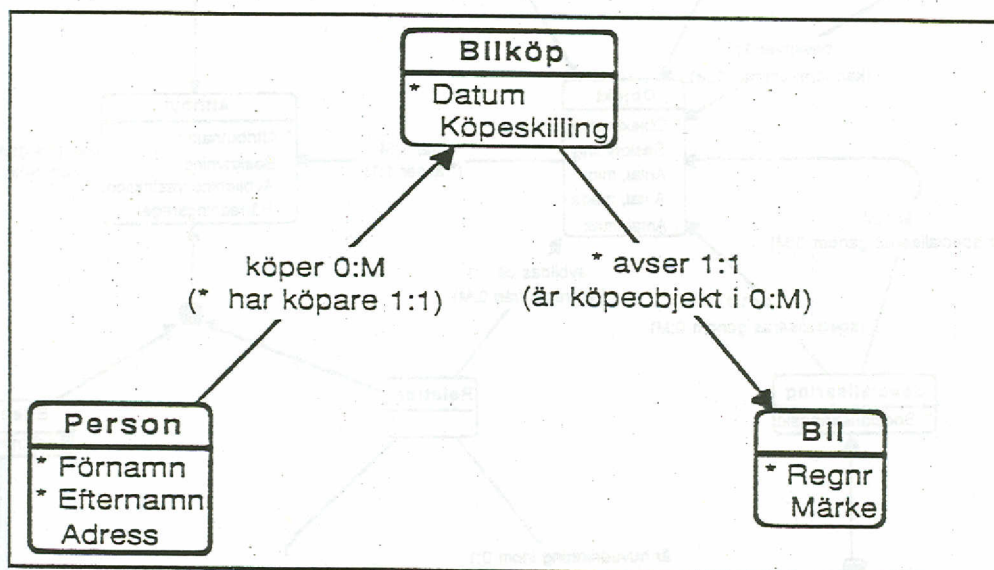


Hur bör en relationsmodell struktureras för att kunna upprätthålla intressanta kopplingar till den semantiskt rikare strukturen för begreppsmodeller, dvs den modell som den genererade relationsmodellen baseras på?

Vi utgår från de grova transformeringsregler mellan begreppsmodell och relationsmodell som redovisas i rapportens två sista sidor samt från följande förutsättningar:

- Med begreppsmodell menar vi här den del av en total begreppsmodell som avgränsats för att svara mot en logisk eller fysisk databas. En logisk databas kan bli horisontellt och eller vertikalt partitionerad av prestandaskäl eller för eventuell fysisk distribution över flera installationer. Både den logiska och den fysiska databasen i form av relationsmodeller antas alltså vara beskrivbara utifrån en avgränsning eller vy av den totala begreppsmodellen.
- En relation representeras av en tabell bestående av ett antal kolumner. I fortsättningen kallas den "tabell".
- En kolumn är enkel, dvs kan inte vara sammansatt av andra kolumner.
- Varje kolumn måste ha ett unikt namn inom tabellen.
- Kolumnordningen saknar betydelse, dvs ordningen kan inte på något sätt användas till att tolka innebörden av en kolumn.

Antag följande begreppsmodell (BM) kring hantering av bilägande.



Transformering i enlighet med redovisad algoritm till relationsmodell (RM) ger upphov till tre tabeller: Bil, Person och Bilköp.

Bil är enklast, med följande uppbyggnad:

Bil	Bil
Regnr (id)	Märke
Märke	Regnr (id)
(alt a)	(alt b)

Om namnet på tabellen sätts till namnet på motsvarande objekt i BM uppstår inga tolkningsproblem vad gäller tabellens innebörd. Det är heller inte svårt att se varifrån kolumnerna i (alt a) härör från i BM. Det blir inte svårare om vi byter plats på kolumnerna, som i (alt b). Deras namn refererar ändå klart till motsvarande begrepp i BM.

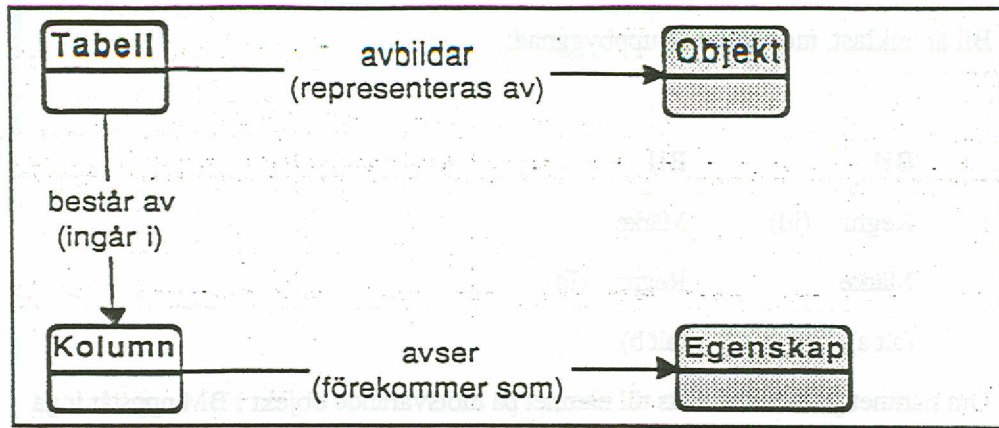
Om alltså både tabell och dess kolumner namnges i enlighet med BM har man dels en implicit koppling till BM, dels möjlighet att därifrån ta fram vilken eller vilka kolumner som bygger upp primärnyckeln (liksom eventuella alternativa nycklar).

Även person-objektet avbildas till en tabell som identifieras med två attribut respektive kolumner. Avbildningen ger sannolikt en tabell enligt (alt a).

Person	Person	Ägare
Förmamn (id)	Efternamn (id)	Tilltalsnamn (id)
Efternamn (id)	Adress	Bostad
Adress	Förmamn (id)	Familjenamn (id)
(alt a)	(alt b)	(alt c)

(Alt b) är likvärdigt eftersom namnen på kolumnerna är desamma som motsvarande egenskaper i begreppsmodellen.

Om man av någon anledning önskar använda andra namn för kolumner, som i (alt c), upphör den namnmässiga kopplingen med begreppsmodellen. Obs, att det visserligen finns anledning arbeta med så stor namnöverensstämmelse som möjligt mellan RM och BM men att det kan finnas fog för olika namn (vilket kommer att visa sig i kommande exempel). Avbildningsprinciperna och kopplingarna mellan RM och BM måste upprättas utan tanke på eventuell namnöverensstämmelse. Konsekvensen blir att både tabeller och kolumner i RM explicit måste peka ut sina respektive ursprung i BM. Se beskrivningsmodellen (metamodellen) nedan utgående från givna förutsättningar. Vilka kolumner som tillsammans utgör tabell-identifieringen, dvs primärnyckel, kan tills vidare härledas via BMs beskrivning av vilka egenskaper som ingår som identifierare.

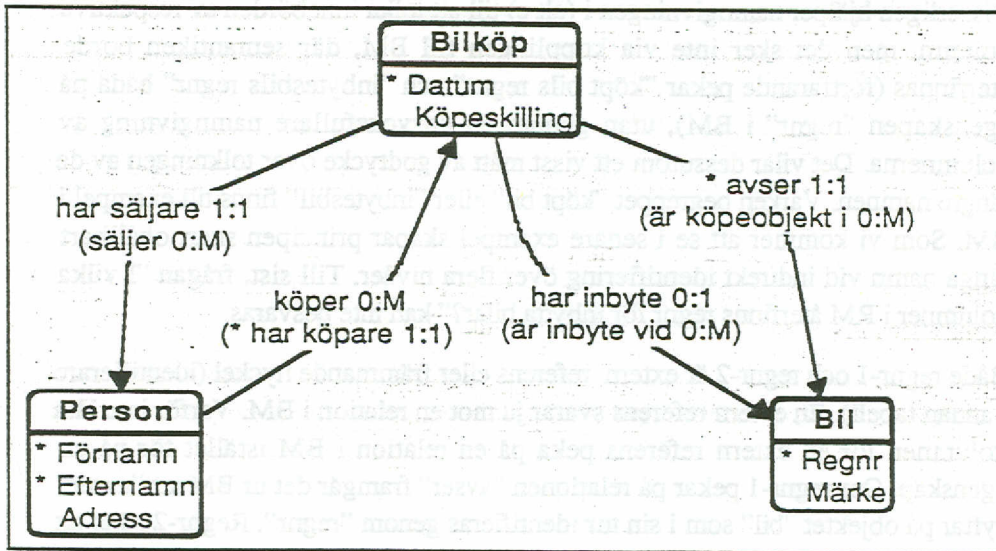


Den tredje tabellen, bilköp, får en uppbyggnad enligt (alt a), om man inte ändrar i avbildningsalgoritmens förslag. Eftersom många omständigheter (namngivningsregler på RM-nivå, prestanda ...) kan motivera förändringar av förslaget måste även en tabell enligt ex.vis (alt b) klaras av.

Bilköp		Ägarskap	
Förmamn	(id)	Ikraftträdande	(id)
Efternamn	(id)	E-namn	(id)
Datum	(id)	Registreringsnummer	(id)
Regnr	(id)	Belopp	
Köpeskilling		F-namn	(id)
(alt a)		(alt b)	

Det nya i denna tabell är att delar av identifieringen hämtas över relationer i BM från andra objekts identifikationer. Med den givna metamodellen kan man fortfarande med viss möda härleda deras motsvarigheter i BM. Kolumnen "f-namn" pekar på person-egenskapen "förmamn". Person är i sin tur relaterad via "köper" med "bilköp". Alltså är "f-namn" förmamnet på den person som är köparen vid bilköpet.

Utvidgar vi exemplet något fungerar inte längre härledningsstrategin. Ta följande utvidgning av bilköpet:



Alternativa motsvarande tabell-utformningar:

Bilköp		Bilköp		Bilköp	
Förmamn	(id)	Förmamn-1	(id)	Köparens förmamn	(id)
Efternamn	(id)	Efternamn-1	(id)	Köparens efternamn	(id)
Datum	(id)	Datum	(id)	Datum	(id)
Regnr	(id)	Regnr-1	(id)	Köpt bils regnr	(id)
Köpeskilling		Köpeskilling		Köpeskilling	
Förmamn		Förmamn-2		Säljarens förmamn	
Efternamn		Efternamn-2		Säljarens efternamn	
Regnr		Regnr-2		Inbytesbils regnr	
(alt a)		(alt b)		(alt c)	

(Alt a) är en direktöversättning av egenskapsnamnen, men strider mot förutsättning d ovan om unika kolumnnamn. (Alt b) har tillförts suffix för att klara av namnkravet och är en korrekt tabell. Men nu tillstöter problem. Både regnr-1 och regnr-2 pekar ut samma egenskap "regnr" i BM. Inget säger att

regnr-1 hör till just den köpta bilen och inte till inbytesbilen. Det kunde lika gärna vara regnr-2, som då skulle noteras som delidentifierare istället för regnr-1.

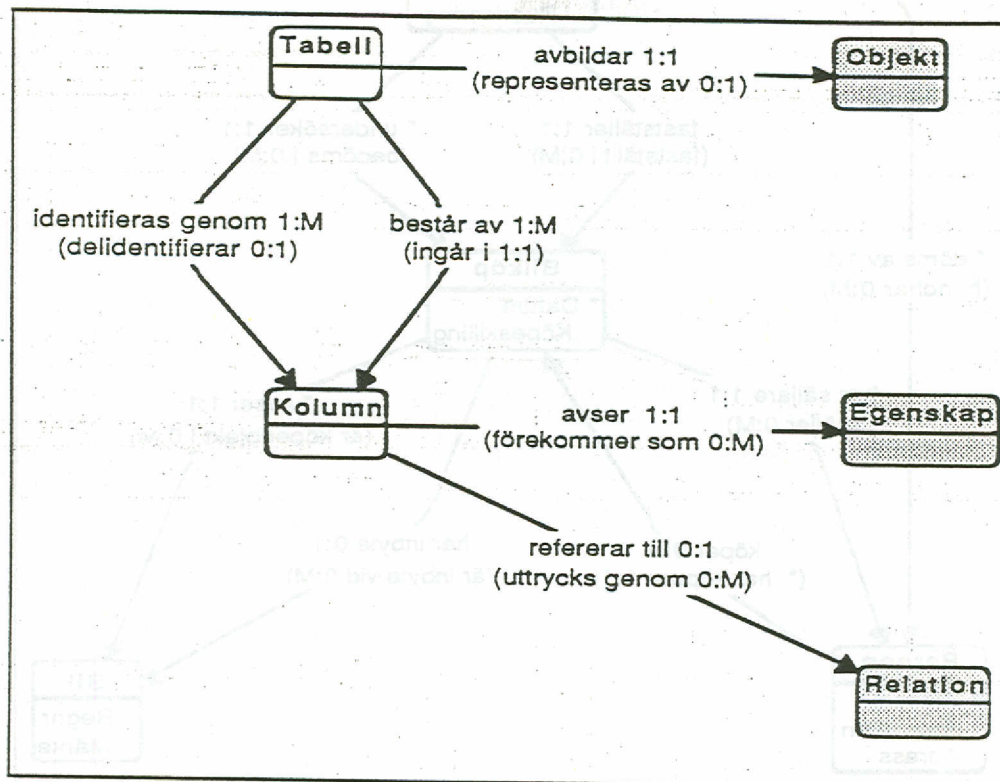
Visserligen hjälper namngivningen i (alt c) till att tolka innebörden av respektive kolumn, men det sker inte via kopplingen till BM, där semantiken borde återfinnas (fortfarande pekar "köpt bils regnr" och "inbytesbils regnr" båda på egenskapen "regnr" i BM), utan genom en uttrycksfullare namngivning av kolumnerna. Det vilar dessutom ett visst mått av godtycke över tolkningen av de längre namnen. Varken begreppet "köpt bil" eller "inbytesbil" finns till exempel i BM. Som vi kommer att se i senare exempel skapar principen snart ohållbart långa namn vid indirekt identifiering över flera nivåer. Till sist, frågan "I vilka kolumner i RM återfinns regnr för inbytta bilar?" kan inte besvaras.

Både regnr-1 och regnr-2 är extern referens eller främmande nyckel (identifierare i annan tabell). En extern referens svarar ju mot en relation i BM. Varför inte låta kolumnen för en extern referens peka på en relation i BM istället för på en egenskap. Om regnr-1 pekar på relationen "avser" framgår det ur BM att "avser" syftar på objektet "bil" som i sin tur identifieras genom "regnr". Regnr-2 sätts att peka på "har inbyte" och entydighet har uppnåtts. Vilka kolumner som deltar i identifieringen av en tabell kan nu också entydigt härledas ur BM. Som komplement kan metamodellen tillföras sambandet "identifieras genom", som direkt pekar ut dessa kolumner.

Är samma lösning överförbar på bilköpets köpare och säljare? Ja, till viss del. Även här gäller det att peka ut rätt relation i BM istället för egenskaper. Skillnaden ligger i att den externa nyckeln består av två kolumner i RM. Låter vi förmamn-1 och eftermamn-1 båda peka på relationen "har köpare" är det entydigt att dessa två tillsammans identifierar den köpande personen. Risker för sammanblandning med förmamn-2 eller eftermamn-2 existerar inte. Vad som dock inte framgår är att förmamn-1 står för personens egenskap "förmamn" och eftermamn-1 för "eftermamn". Observera, återigen, att vi i resonemanget måste bortse från att namnen på kolumn och egenskap är snarlika. Vi måste dels tala om vilken egenskap kolumnen motsvarar, dels vilken relation de båda kolumnerna tillsammans svarar mot.

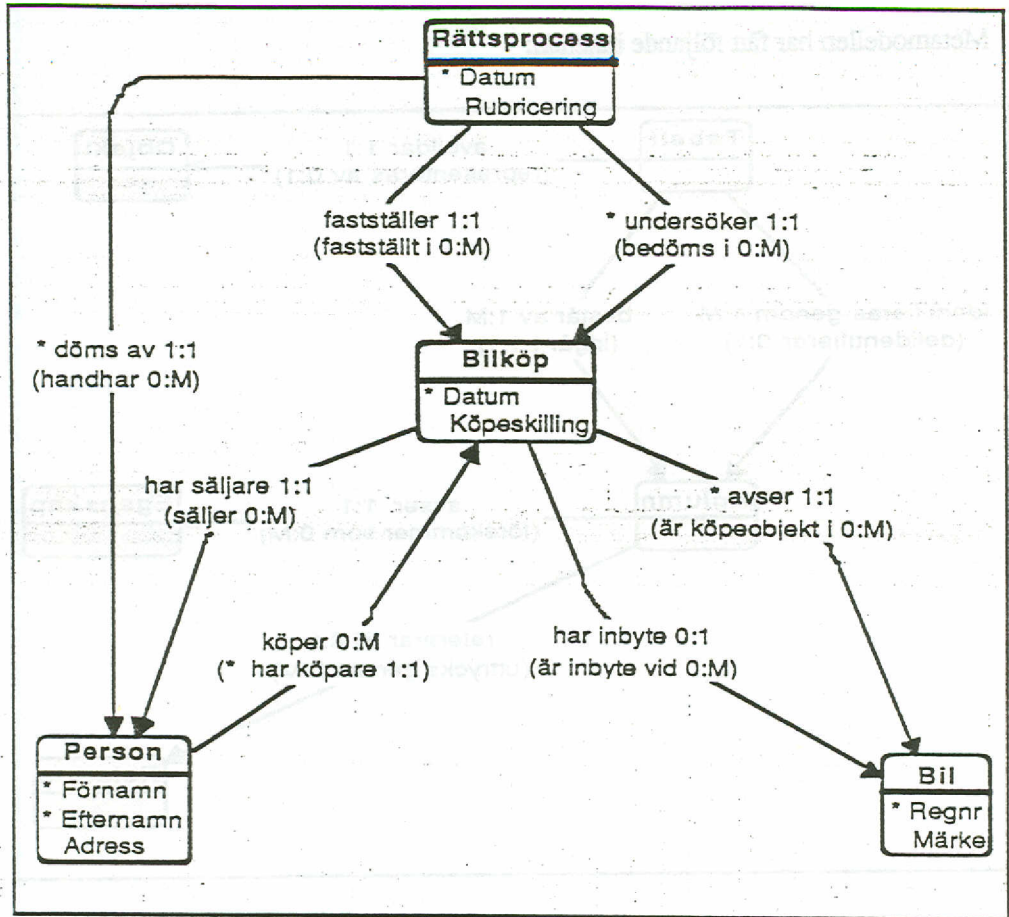
Det blir ingen synlig ändring i metamodellen, bara en vetskap om att flera kolumner ur samma tabell nu kan referera till samma relation. Därtill kommer att en och samma kolumn ibland pekar på både en egenskap och en relation.

Metamodellen har fått följande innehåll:



Det som nu utretts är referens mellan egenskaper och kolumner samt mellan relationer på en nivå och deras kolumnmotsvarigheter. Är metamodellen därmed även generellt heltäckande?

Vi testar genom att komplettera begreppsmodellen enligt nedan. Exemplet kan måhända upplevas som konstruerat men saknar inte realism. Denna typ av situation liksom ännu mer komplexa dito måste kunna hanteras i en metamodell.



Objektet "rättsprocess" ger upphov till en av följande tre alternativa tabeller:

Rättsprocess	Rättsprocess	Rättsprocess
Förnamn-1 (id)	Köparens förnamn (id)	Undersökt ägarskaps köparens förnamn (id)
Efternamn-1 (id)	Köparens efternamn (id)	Undersökt ägarskaps köparens efternamn (id)
Datum-1 (id)	Bilköpsdatum (id)	Undersökt ägarskaps datum (id)
Regnr-1 (id)	Köpt bils regnr (id)	Undersökt ägarskaps bils regnr (id)
Förnamn-2 (id)	Domarens förnamn (id)	Domarens förnamn (id)
Efternamn-2 (id)	Domarens efternamn (id)	Domarens efternamn (id)
Datum (id)	Rättsprocessdatum (id)	Rättsprocessdatum (id)
Förnamn-3	Köparens förnamn-2	Fastställt ägarskaps köparens förnamn
Efternamn-3	Köparens efternamn-2	Fastställt ägarskaps köparens efternamn
Datum-3	Bilköpsdatum-2	Fastställt ägarskaps datum
Regnr-3	Köpt bils regnr-2	Fastställt ägarskaps bils regnr
Rubricering	Rubricering	Rubricering

(alt a) (alt b) (alt c)

(Alt a) ger den torftigaste förståelsen. Vi har exempelvis så mycket som tre stycken förmamn som vi inte vet exakt vad de står för. Även om vi antar att (id)-angivelserna hamnat rätt finns fortfarande två förmamn kvar inom identifieringen, vars roller inte framgår. (Alt b) svarar ungefär mot den semantik som är uttryckbar genom metamodellen.

I det tidigare exemplet lät vi för- och efternamnen, förutom på egenskaper i BM, även peka på den relation som motiverat deras respektive medverkan i relationen. En kombination för- och efternamn pekade på "har köpare", den andra på "har säljare". Nu räcker inte den lösningen längre. Vi har två kolumner som båda pekar på egenskapen "förmamn" och på relationen "har köpare". Samma sak gäller för efternamn. Tillsammans pekar alltså fyra kolumner på "har köpare".

Orsaken finns att hämta i de båda roller "bilköp" befinner sig under "rättsprocess", nämligen både som undersökt och som fastställt bilköp. I båda fallen del-identifieras bilköpet med identifieraren för köparen. Det måste kunna gå att uttrycka vilket par av för- och efternamn som hänvisar till köparen i det undersökta bilköpet och vilket par som gäller för köparen i det fastställda bilköpet. (Alt c) ovan visar en semantik i kolumn-benämningen som i praktiken måste kunna framgå genom referenserna till BM. Vi har redan tidigare konstaterat att det inte går att förlita sig på namngivningen av kolumner. Det kan t.ex. säkert finnas situationer som är betydligt mer komplexa än denna.

Entydighet kan bara uppnås om varje relation kopplas till den kombination kolumner som svarar mot den främmande nyckel relationen uttrycks genom i en tabell. Vi måste exempelvis skapa den grupp kolumner som utgör den främmande nyckeln för bilköp i rollen av undersökt bilköp. Tjuvtittar vi på alternativet c men väljer benämningarna enligt alternativ a (för att ta det mest intetsägande alternativet) får vi gruppen:

förmamn-1

efternamn-1

regnr-1

datum-1

men det skulle lika gärna kunnat vara (men då med annan placering av id-noteringar):

förmamn-3

efternamn-2

regnr-1

datum

Motsvarande gruppering måste göras för fastställt bilköp. Vi kan låta fyra kolumner peka på relationen "fastställer" och fyra andra på relationen "undersöker" enligt tidigare strategi. Dock finns inom respektive gruppering två andra främmande nycklar, nämligen den som pekar ut person i rollen som köpare och bil i rollen av köpt objekt. Ska ex.vis kolumnen förmamn-1 peka både på "undersöker" och "har köpare"? Här räcker det med "undersöker" eftersom det i detta exempel inte finns någon risk att inom den större grupperingen missförstå betydelsen av förmamn. Vi har bara ett förmamn inom gruppen. Det kan entydigt härledas till köparen.

Därmed inte sagt att det inte kan finnas situationer där flera förmamn uppträder inom gruppen liksom del-grupperingar med samma egenskapsreferenser. Exempel på detta är grupperingen av de kolumner som utgör identifieraren för en rättsprocess. Denna kan ju tänkas ingå som delidentifierare i något annat sammanhang. Semantiken måste kunna framgå entydigt även i de mest komplicerade fallen. Nyckeln ser ut som följer:

Identifierare/primärnyckel

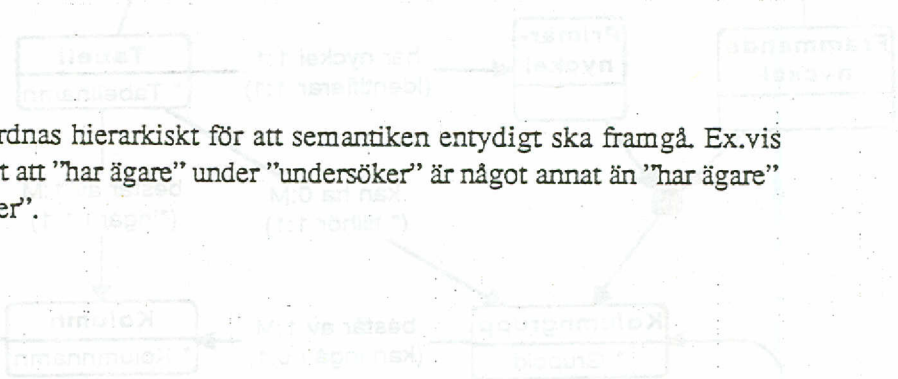
för rättsprocess	undersöker	har ägare	avser döms av
förmamn-1	förmamn-1	förmamn-1	
eftermamn-1	eftermamn-1	eftermamn-1	
regnr-1	regnr-1		regnr-1
datum-1	datum-1		
förmamn-2			förmamn-2
eftermamn-2			eftermamn-2
datum			

Där nyckeln figurerar som en främmande nyckel räcker det inte att låta dess kolumner peka på den relation som givit upphov till den främmande nyckeln. Vi måste inom nyckeln skilja ut ex.vis rollen för förmamn-1 från rollen för förmamn-2. Subgruppering av referenser till "undersöker" respektive "döms av" blir nödvändig. Resultatet blir att förmamn-1 ingår i primärnyckeln för rättsprocess och i sub-gruppen som pekar på "undersöker". Förmamn-2 ingår i samma primärnyckel samt i sub-gruppen som pekar på "döms av". För att slippa arbeta med härledning (där detta går) väljer vi att konsekvent skapa en kolumngrupp för varje främmande nyckel. Kolumngruppen för "har ägare" under nyckeln skulle ex.vis för närvarande inte behövas, men vad säger att denna situation för alltid består?

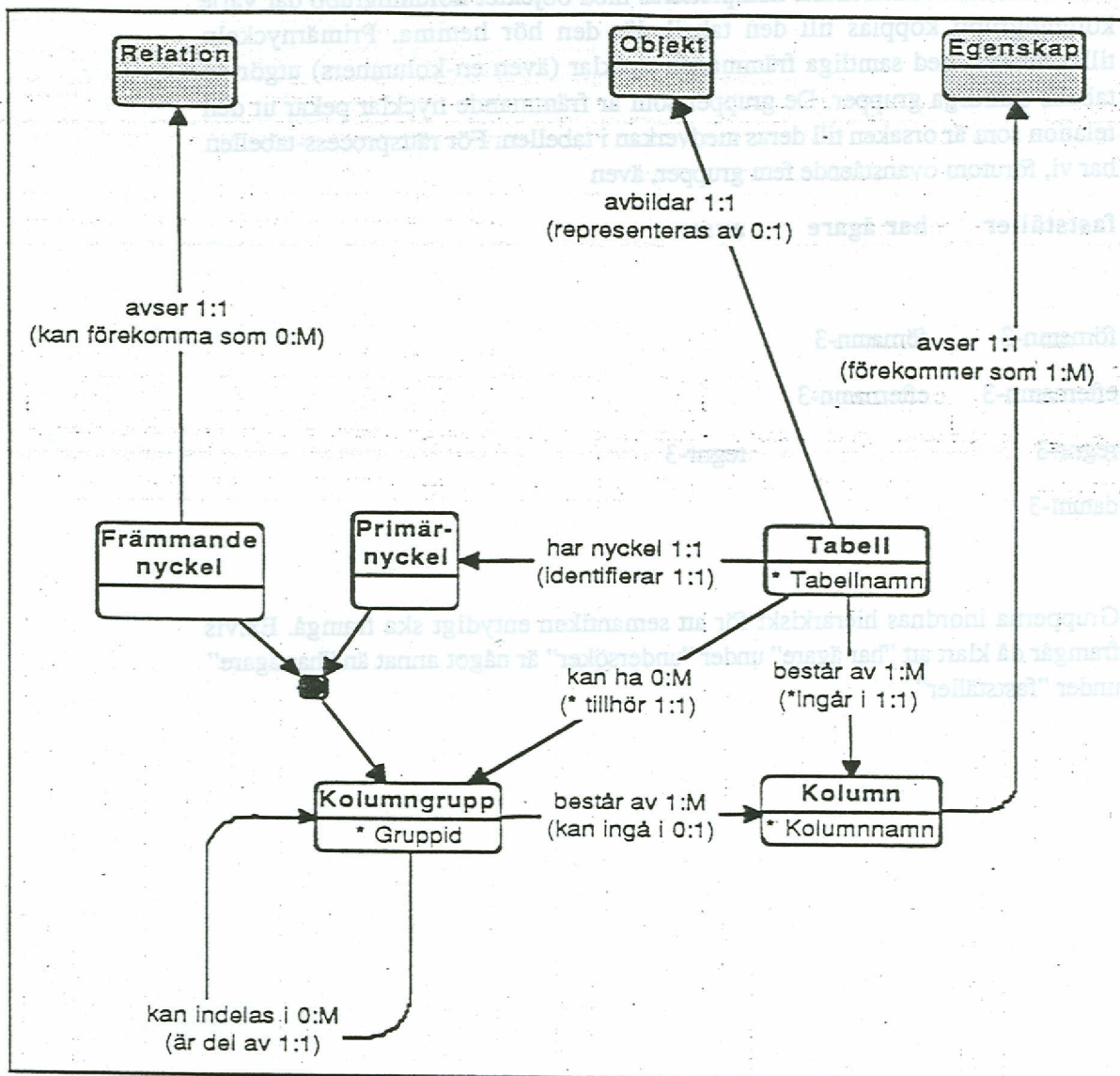
M.a.o. metamodellen måste kompletteras med objektet kolumngrupp där varje kolumngrupp kopplas till den tabell där den hör hemma. Primäryckeln tillsammans med samtliga främmande nycklar (även en-kolumners) utgör en tabells samtliga grupper. De grupper som är främmande nycklar pekar ut den relation som är orsaken till deras medverkan i tabellen. För rättsprocess-tabellen har vi, förutom ovanstående fem grupper, även

fastställer har ägare avser
 förmamn-3 förmamn-3
 eftermamn-3 eftermamn-3
 regnr-3 regnr-3
 datum-3

Grupperna inordnas hierarkiskt för att semantiken entydigt ska framgå. Ex.vis framgår då klart att "har ägare" under "undersöker" är något annat än "har ägare" under "fastställer".

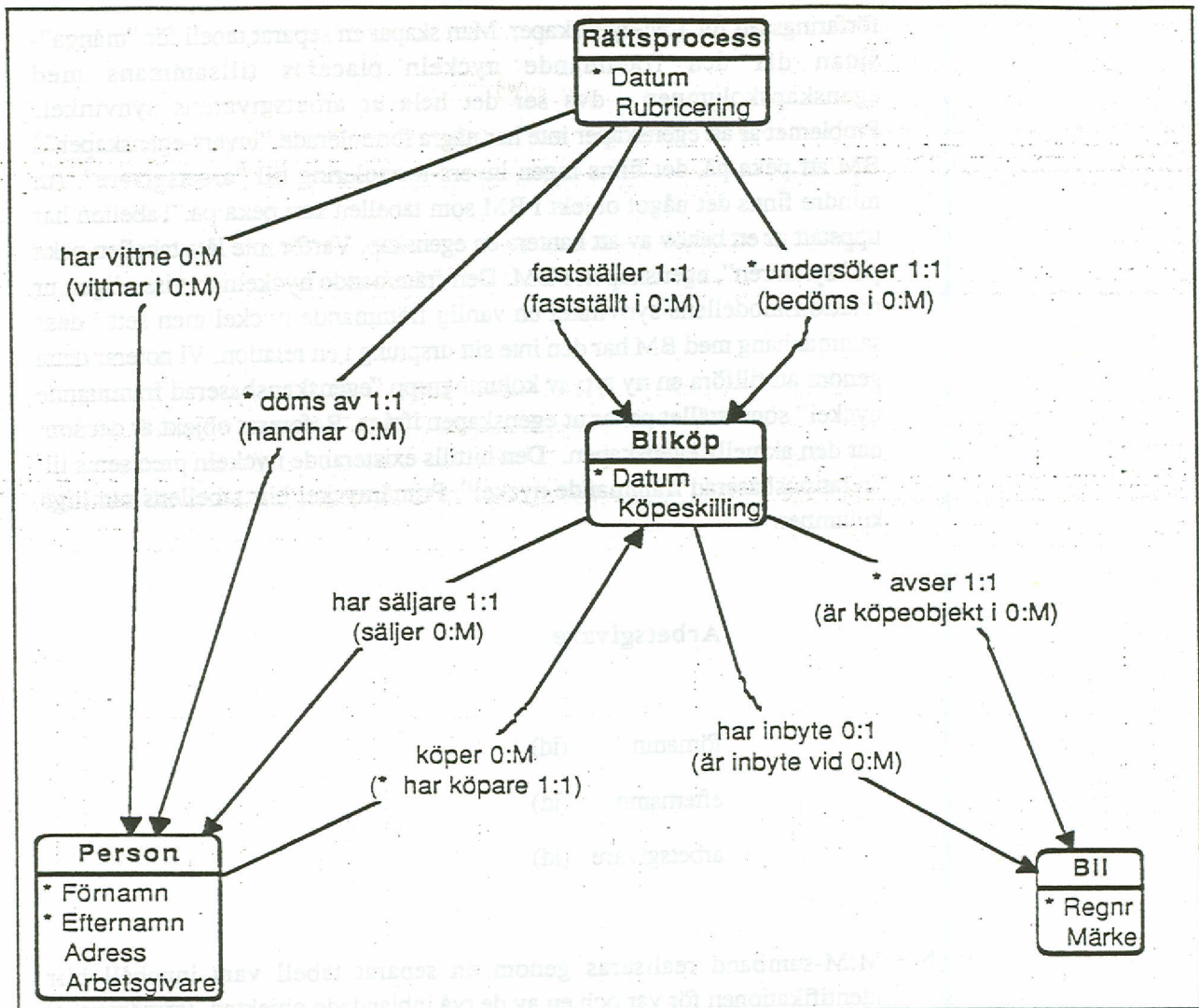


Metamodellen presenteras nedan i sitt kompletterade skick.



Till sist belyser vi hur de tidigare presenterade transformeringsreglerna påverkar kopplingen mellan BM och RM. Inkluderat i den existerande rättsprocessmodellen är 1:M-samband (ex.vis över relationen "köper"), M:1-samband (ex.vis över relationen "undersöker") och envärda egenskaper. 1:1-samband kräver bedömning kring var i de två tabellerna den främmande nyckeln ska placeras men hanteras annars i princip enligt M:1-samband. Kvår att behandla är flervärda egenskaper och M:M-samband.

För att beskriva dessa två varianter kompletteras rättsprocessmodellen enligt följande:



Tillfogad är följande:

- a. Egenskapen "arbetsgivare" till objektet person. En person kan ha flera arbetsgivare.
 - b. De personer som inkallas som vittnen under processen. En process kan ha flera vittnen. En person kan vara vitne vid flera processer.
- a. Väljer man att stanna på 1:a normalformen blir det inga andra ändringar än att tabellen för person kompletteras med en kolumn för arbetsgivare. Kolumnen pekar ut egenskapen "arbetsgivare" i BM.

Om målet är tabeller enligt 3:e normalformen behöver denna 1:M-egenskap hanteras speciellt. Vid 1:M-samband placerar vi enkelt den främmande nyckeln på "många"-sidan, dvs med M:1-aspekten, samt sätter den att peka på sambandets M:1-relation i BM. I grunden skulle man önska samma

förfaringssätt för 1:M-egenskaper. Man skapar en separat tabell för "många"-sidan där den främmande nyckeln placeras tillsammans med egenskapskolumnen, dvs ser det hela ur arbetsgivarens synvinkel. Problemet är att egenskaper inte har några formulerade "invers-egenskaper" i BM att peka på, det finns ingen invers-formulering till "arbetsgivare". Än mindre finns det något objekt i BM som tabellen kan peka på. Tabellen har uppstått ur ett behov av att hantera en egenskap. Varför inte låta tabellen peka på "syndaren", egenskapen i BM. Den främmande nyckeln är visserligen ur relationsmodellens synvinkel en vanlig främmande nyckel men sett i dess sammanhang med BM har den inte sitt ursprung i en relation. Vi noterar detta genom att tillföra en ny typ av kolumngrupp "egenskapsbaserad främmande nyckel" som istället pekar ut egenskapen ifråga. Refererat objekt är det som har den aktuella egenskapen. Den hittills existerande nyckeln preciseras till "relationsbaserad främmande nyckel". Primärnyckel blir tabellens samtliga kolumner.

Arbetsgivare

förmamn (id)

eftermamn (id)

arbetsgivare (id)

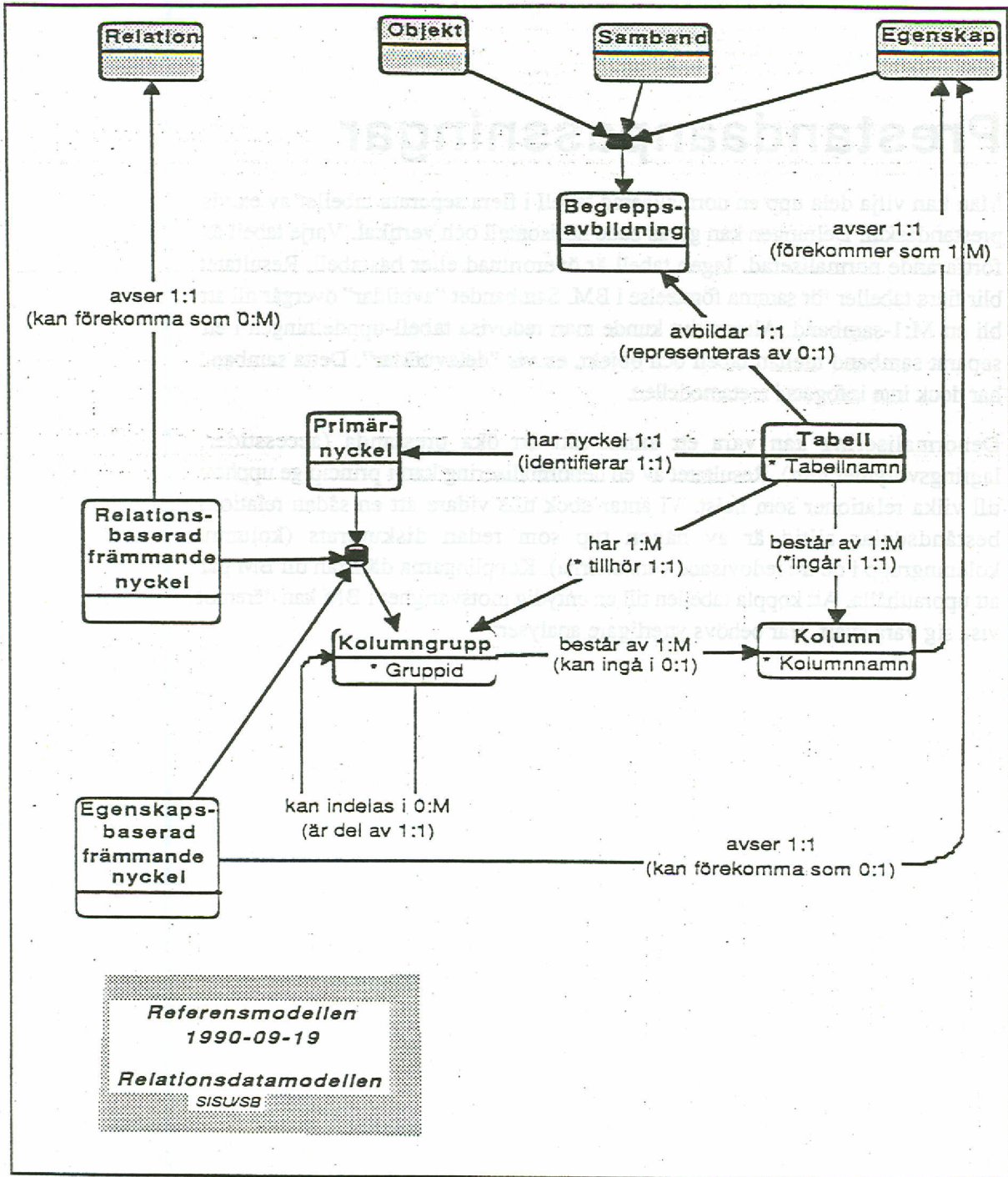
- b. M:M-samband realiseras genom en separat tabell vars innehåll blir identifikationen för var och en av de två inblandade objekten. Primärnyckel är summan av dessa två kolumngrupper. Var och en av de två kolumngrupperna är en främmande nyckel som sätts att peka på den relation som står för ett M:1-förhållande till objektet. I exemplet pekar den främmande nyckeln för "rättsprocess" på "vittnar i" medan nyckeln för "person" pekar på "har vittne". Det är lättare att se situationen om man upplever tabellen som en representation av "person som vittne i viss rättsprocess". I dennes specifika perspektiv finns ju bara en person, som återfinns över "har vittne" och en rättsprocess, som återfinns över "vittnar i".

Den enda komplikationen uppstår vid relateringen av tabellen till dess motsvarighet i BM. Ursprunget är inte ett objekt utan ett samband.

Prestandaanpassningar

Man kan vilja dela upp en normaliserad tabell i flera separata tabeller av ex.vis prestandaskäl. Delningen kan göras både horisontell och vertikal. Varje tabell är fortfarande normaliserad. Ingen tabell är överordnad eller bastabell. Resultatet blir flera tabeller för samma företeelse i BM. Sambandet "avbildar" övergår till att bli ett M:1-samband. Alternativt kunde man redovisa tabell-uppdelningen i ett separat samband mellan tabell och objekt, ex.vis "delavbildar". Detta samband har dock inte infogats i metamodellen.

Denormalisering kan vara ett annat sätt att öka prestanda (accesstider, lagringsvolym mm). Resultatet av en denormalisering kan i princip ge upphov till vilka relationer som helst. Vi antar dock tills vidare att en sådan relations beståndsdelar alltid är av någon typ som redan diskuterats (kolumn, kolumngrupp i de tre redovisade versionerna). Kopplingarna därifrån till BM går att upprätthålla. Att koppla tabellen till en entydig motsvarighet i BM kan däremot visa sig vara svårt. Här behövs ytterligare analyser.



Grova tumregler för transformation från begreppsmodell till relationsmodellen enligt 3:e normalform.

Denna kortfattade utredning har gjorts med syftet att utröna vilka kopplingar mellan begreppsmodell och relationsmodell som kan bli aktuella att avbilda i referensmodellen.

Egenskapsfokusering

- a. För varje objekt som har envärda egenskaper (E), skapa tabell med objektets primärnyckel (OPN) samt de envärda egenskaperna. (OPN, E1,...,En)

Ev. optimering: Undvik många NULLs genom att bryta ut NULL-rika egenskaper (eller egenskapsgrupper) till egna tabeller. (OPN, E1,...,Ex),..., (OPN, Ey),..., (OPN, Ez,...En)

- b. För varje flervärd egenskap, skapa tabell för varje sådan egenskap innehållande objektets primärnyckel och egenskapen. (OPN,E)

Sambandsfokusering

- c. För varje 1:1-samband av typen

0,1:0,1 placera objektnyckeln för objektet med störst kardinalitet i tabellen för objektet med minst kardinalitet som främmande nyckel.

$O1 \text{ ---} 0,1 \text{ ---} s \text{ ---} 0,1 \text{ ---} O2$ blir $O1: (OPN1, OPN2, \dots)$
 $O2: (OPN2, \dots)$

100 100000

Ev. optimering: Om många NULLS i referensen, placera sambandet i separat tabell. (OPN1,...), (OPN2,...), (OPN1, OPN2)

0,1:1,1 placera främmande nyckeln i tabellen för objektet med 1:1-koppling för att undvika NULLS.

$O1 \text{ ---} 0,1 \text{ ---} s \text{ ---} 1,1 \text{ ---} O2$ blir $O1: (OPN1, \dots)$
 $O2: (OPN2, OPN1, \dots)$

1,1:1,1 välj endera objekt-tabellen.

$O1: (OPN1, OPN2, \dots)$ och $O2: (OPN2, \dots)$

eller

$O1: (OPN1, \dots)$ och $O2: (OPN2, OPN1, \dots)$

- d. För varje 1:M- eller M:1-samband, placera nyckeln på tabellen för objektet på M-sidan.

$O1 \text{---}x,1 \text{---}s \text{---}x,M \text{---} O2$ blir $O1: (OPN1, \dots)$
 $O2: (OPN2, OPN1, \dots)$

($x=0$ eller 1)

Ev. optimering: Om många NULLS i främmande nyckeln, placera sambandet i separat tabell.

$O1 \text{---}0,1 \text{---}s \text{---}x,M \text{---} O2$ blir $(OPN1, \dots), (OPN2, \dots)$
 $(OPN1, OPN2)$

($x=0$ eller 1)

- e. För varje M:M-samband skapas en separat tabell med båda objektens nycklar.

$O1 \text{---}x,M \text{---}s \text{---}x,M \text{---} O2$ blir $(OPN1, OPN2)$

Specialiseringsfokusering

- f. För super/sub-samband om

super har id	se i princip 1:1-samband ovan där åtminstone 0,1(super):1,1(sub) gäller. Har sub eget id används detta annars ärvs supers id.
inget superid	implicerar olika id på subnivå. Som ovan, men först måste artificiell id skapas för super.

Optimeringar därutöver

är beroende av

- vald teknisk lösning
- tänkta framtida anpassningar och utvidgningar av modell
- systemsamordning

mm

och behandlas därför inte vidare i detta forum.

TRIAD utvecklar IA

Televerket har just tagit första steget in i sin nya IA-organisation och Posten håller på att bygga upp sin nya DA-organisation. Båda organisationerna har sett nytan att inför 90-talet gå vidare tillsammans i TRIAD-projektet som drivs tillsammans med SISU. Statskontoret deltar också i projektet för att på sikt kunna föra ut nya synsätt och hjälpmedel inom den civila statliga sektorn.

Ericsson Data Services deltar med tyngdpunkten i den del som handlar om att utveckla kompetenta modelleringsledare, delprojektet "Avancerad utbildning för modelleringsledare".

Modelleringsmetoder är centrala i bedrivandet av verksamheten inom informationsadministrationen. Därför arbetar ett delprojekt med utvecklandet av "nästa generation modelleringsmetod" som skall sättas i händerna på informationsadministratören. Siktet är att fördjupa och bredda dagens modelleringsmetoder och där hämta in kunskap från pågående forskning och utveckling internationellt. (faktaruta om IAS91).

Som stöd för informationsadministrationen behövs verktyg. Inom TRIAD arbetar man där inom två områden, kataloger och verktyg.

Delprojektet kataloger arbetar dels med att utforma den informationsmodell som måste kunna täckas av en katalog, dels med att granska och följa utvecklingen av produkter inom området t ex IBM:s "Repository" och Digital's "CDD". Dessutom följer man standardiseringen internationellt kring IRDS. För parterna i projektet liksom för andra organisationer är detta ett tungt område både vad gäller kommande investeringar ekonomiskt och vad gäller kompetenta resurser för en kommande övergång till "repository-världen". - Det inledande skedet syftar till att bygga upp en kunskapsplattform, som sedan kommer att kunna utnyttjas för kravställande och planering och genomförande av övergång från dagens kataloghantering till morgondagens.

Den andra verktygshanterande delen inom TRIAD-projektet, delprojektet "verktyg för informationsadministration", syftar till att ta fram verktyg för uttag och dokumentering av modeller. Betoningen ligger på människa datorgränssnitt och i första skedet görs utveckling av HYBRIS-gränssnittet med prototyper för Posten och för Televerket.

För att hålla ett helhetsperspektiv på projektets delar och för att ha inpassningen av funktionen Informationsadministration i organisationens övriga verksamhet arbetar delprojektet "Krav på IA". I delprojektet arbetar man dels med att kartlägga dagens krav på dataadministration och projicera till morgondagens krav på IA. Dessutom skall man skapa en bild av IA-verksamhetens innehåll och organisation. Från detta i sin tur ställer man krav

på övriga delprojekt. Vilka krav skall ställas på kompetens, metoder, hjälpmedel typ kataloger och gränssnitt?

TRIAD projektet är stort

Budgeten för TRIAD-projektet löper på 10 MSEK per år under en treårsperiod som startar vid kalenderåret 1991 års början och som alltså beräknas avslutad vid utgången av 1993.

TRIAD-projektet är ett tillämpningsprojekt

Det innebär att parterna, Televerket, Posten, Statskontoret, EDS och SISU går in med såväl persontidsansatningar som ekonomiska och att STU, Styrelsen för Teknisk Utveckling, bidrar med ett ekonomiskt tillskott som svarar mot ungefär 40 % av den insatta persontiden.

Öppet för fler deltagare

Parterna i TRIAD-projektet vill gärna öka tempot och bredda perspektivet och vill därför gärna ha fler parter in i projektet. Dessa parter får då enligt SISU:s tårtprincip "betala för en tårbit, men ät hela tårtan", tillgång till projektets resultat med en insats som ger stor "price performance".

Nya deltagare kan gå in i hela projektet eller i det eller de delprojekt som verkar intressantast. En förutsättning är att man framförallt är beredd att satsa kompetent personal. För de flesta intressenter bord detta vara ett utmärkt sätt att driva personalutveckling för personer t ex inom DA-området, samtidigt som man bygger upp beredskapen inför 90-talets IA-verksamhet.

Kompetensutveckling viktigt resultat

En viktig effekt för parterna av deras medverkan i TRIAD är kompetensutveckling. Man satsar på att ta in personer som så småningom eller redan idag arbetar med DA och IA för att ge dem en djup och "frontlinje"-mässig kompetens. Detta skall utnyttjas när man successivt för in resultaten i den egna organisationen. Projektdeltagarna har alltså en viktig roll som kunskapsförmedlare i den egna organisationen. Dessutom ger projektarbetet deltagarna tillfälle till en egen utveckling inom det professionella området som är unik.

Informationsspridning

Det sjätte delprojektet "Informationsspridning" har till uppgift att sörja för att i första hand parterna men också SISU:s övriga intressenter successivt kan följa och tillgodogöra sig resultat från TRIADprojektet. Seminarier, rapporter och referensgruppsverksamhet är led i den verksamheten.