

Tesi doctoral presentada per En/Na

Elisabet ROSA TRÍAS

amb el títol

**"Conceptes, tècniques i problemes de la
generalització cartogràfica"**

per a l'obtenció del títol de Doctor/a en

GEOGRAFIA

Barcelona, 25 de febrer del 2000.

Facultat de Geografia i Història
Departament de Geografia Física i Anàlisi Geogràfica Regional



UNIVERSITAT DE BARCELONA



3. L'escala en la compilació de les dades espacials

El cambio de escala corresponde a un cambio del nivel de análisis y debería corresponder a un cambio del nivel de conceptualización.
(Lacoste 1977:55)

El pas, en el tractament de les dades espacials, d'un entorn manual a un de digital, ha suposat el replanteig tant dels antics processos com dels seus objectius en el nou entorn. El procés de generalització, part integrant d'un procés més ampli de tractament de la informació espacial, també s'ha vist obligat a haver de canviar les tasques per adaptar-se a l'entorn digital, no pas del procés de generalització cartogràfic manual, sinó revisant-ne els objectius en el nou entorn. L'entorn digital obliga a formalitzar les operacions per poder automatitzar-les, i això és problemàtic, ja que és un procés difícil d'explicitar i per tant, pel que fa a funcions, també és difícil d'estructurar. La dificultat, des del moment que es comencen a confeccionar els SIG fins a l'actual dècada, ha fet que el procés de generalització s'hagi reduït a les representacions cartogràfiques de les formes geomètriques, sense tenir en compte que sovint representen fenòmens de la superfície terrestre diferents i que per tant no admeten un mateix operador, tot i que la funció pugui ser la mateixa per a diversos fenòmens. Per exemple l'operador o criteri per simplificar una línia de costa no ha de ser el mateix que hem d'utilitzar per simplificar una àrea administrativa, encara que la funció de simplificació sigui comuna als dos fenòmens.

El preprocessament de les dades espacials és essencial en el procés de generalització en un entorn digital, ja que permet a la generalització entrar en un terreny més abstracte: el del tractament de les dades originals, no pas en funció de les necessitats de la comunicació, de les representacions cartogràfiques, sinó en funció de les necessitat de l'anàlisi. Ja hem vist en l'anterior capítol que aquests processos han hagut d'adaptar-se al nou entorn, no solament emulant la generalització en l'entorn manual sinó fent factible un tractament previ de les dades. Aquests passos previs que pertanyen al preprocessament permeten la federació de bases de dades per integració i

homogeneïtzació, sota un control diferent del de l'antic "cop d'ull" del cartògraf experimentat. Aquesta possibilitat de tractament és important per facilitar l'accés a les dades. El seguiment durant aquest preprocessament, l'anomenem visualització de les dades, i no és la finalitat que perseguim, sinó que és només una possibilitat de controlar el procés i d'ajudar en l'anàlisi de les dades en el DLM.¹ La visualització no prové del camp de la cartografia tradicional, sinó que pertany a la tradició de presentació de les dades d'altres disciplines, com ara les que tracten dades de tipus no espacial.

Encara que gran part de les dades de què disposem provenen de la digitalització de mapes i estan conceptualment lligades a les escales a què aquests van ser concebuts (Monmonier 1993:102), en un entorn digital la principal font d'informació geogràfica ja no és el mapa sinó que sobretot la base de dades, on les dades recollides directament al camp, o a través de sensors remots, s'integren en un model digital de dades primari. I en aquestes bases de dades el concepte d'escala geogràfica s'ha perdut.

"A map, and more generally a spatial database, is a statement about geographical reality and generalization is a process which tends to universalize the content of that statement" ([Muller, 1991 #471]:457).

La base de dades, com el mapa, presenten una visió determinada de la realitat, amb l'objectiu de generar nova informació. Actualment es considera que la generalització ha de permetre la derivació de noves dades a partir d'una base de dades prèvia, així com la visualització de les dades per pantalla, que actua com a interfície entre la informació i l'usuari. I respecte això, Muller fa notar que:

"The way the data model is organized and can be generalized is likely to influence the performance of cartographic generalization" (Müller et al. 1995b:5)

Des del punt de vista de l'usuari, la finestra de l'ordinador és el mitjà a través del qual pot manipular les dades per extreure'n informació i, posteriorment, coneixement geogràfic. És, per tant, important que l'associació lògica entre els objectes espacials i el

¹ La visualització és també utilitzada en el camp més general dels sistemes espacials per part de diferents disciplines que utilitzen la representació espacial, encara que no ho facin a les mateixes escales que ho fa la geografia. La visualització de les dades no ha de servir les regles de comunicació de la cartografia, ja que és un pas intermedi en el model de les dades, no basat estrictament en la comunicació, sinó més aviat en les noves teories perceptives i de

seu significat geogràfic es pugui preservar a totes les escales, independentment de les característiques que resumeixin la naturalesa de la majoria de fenòmens geogràfics. Els processos geogràfics són dependents de l'escala, i diferents aplicacions comporten diferents nivells de detall; hi ha, per tant, una necessitat de modelització de la informació geogràfica als diferents nivells d'abstracció, i tal com diuen Tobler & Mollering (1972) i Woodcock & Strahler (1987):

"Ideally one should be able to view and analyze data at the level where geographical variance is maximized" (Müller et al. 1995b:3)

És a dir, hauríem de poder veure i analitzar les dades en el nivell on els processos espacials poder ser més ben compresos, i això és impossible de moment, ja que les bases de dades no contenen prou informació, i no es coneixen ara com ara els processos de generalització per poder derivar d'un model de dades original, de manera automàtica, un altre model a una resolució desitjada.

En el moment actual, qui controla el procés de derivació i decideix sobre el nivell de detall de les característiques mètriques i no mètriques o dels atributs, és el cartògraf. És qui controla el procés de generalització basant-se en el coneixement geogràfic i de les formes, tenint en compte la finalitat del mapa i l'escala a què es vol analitzar les dades o prendre una decisió. Per tant, hem d'aconseguir emular els coneixements del cartògraf, o sigui, fer el procés suficientment "intel·ligent".

Fins fa uns anys, aquesta ha estat la tendència, la de creure que la solució de l'automatització consisteix a fer el procés més intel·ligent, dissenyant algorismes més adequats per a operadors que es corresponguin millor amb el tipus de funció que han de desenvolupar. Però, en l'actualitat, aquest enfocament basat en la millora del procés s'ha vist canviat per la creença que és necessari poder "enriquir" les dades, de manera que continguin saber des de l'inici, per poder guiar el tipus de processament de manera més adequada.

Buttenfield assenyala que seria també un gran ajut disposar d'uns models de dades més adequats, que facilitessin la manera d'organitzar les dades, ja que això influeix en la seva manipulació posterior. L'usuari no hauria de saber com estan estructurades

físicament les dades en la base de dades per poder treballar. Això podria simplificar el procés de codificació de la informació o representació gràfica (*surface structure*) d'un model de dades que contingués les relacions i els atributs espacials en "l'estructura interna" (*deep structure*), ja que tant les operacions de generalització en el model de dades a nivell "intern" com la recollida i la compilació de les dades tenen un efecte en la superfície o nivell "extern" (Buttenfield 1995:91).

L'ampliació del procés de generalització en l'entorn digital pretén afegir com un subprocés amb entitat pròpia allò que pels cartògrafs consistia simplement a "donar un cop d'ull" a la informació en el preprocessament o procés de compilació. El preprocessament en un entorn digital es pot realitzar en el model de les dades per tractament estadístic; per això, ara es disposa d'eines d'estadística espacial. Muller es refereix a aquest preprocessament de les dades espacials quan diu que és necessari que el cartògraf s'involucri en la precartografia.

“...the realm of cartographic activities has expanded towards the portrayal of highly abstract geographical spaces...a stronger involvement of cartographers in precartography (data quality and resolution) will emancipate cartographic thinking by a platonic view of a world made of ideal forms and help design cartographic products what-you-get-you-cannot-see” (Muller 1989:675).

Ruas i Lagrange també veuen la necessitat que el camp de la generalització s'eixampli, en especial al que fa referència al model de les dades, com a pas al model cartogràfic, “several issues related to geogràfic data modelling and GIS content are arising [...] representational needs of cartographic generalization become needs of gis systems themselves”.

La generalització és fonamental en la gestió de la informació i en la preparació de les dades per integrar diferents bases de dades, de resolucions i procedències diferents, que obliguen a fer una que s'homogeneïtzi a una escala d'anàlisi comuna ç

“...it is the right tool for various operations such data integration or deriving of special informació from a databases” (Lagrange, Ruas 1994:1099)

El fet de no considerar la generalització del model de les dades com un procés que es realitza fora de l'entorn del SIG i de considerar la generalització en la seva vessant de

representació cartogràfica, com s'ha fet fins fa poc, amaga els importants efectes que provoca posteriorment en l'anàlisi. Actualment això és més evident, ja que es reconeixen tant les necessitats com l'interès de la generalització en els sistemes d'informació geogràfica. Així, doncs, no considerem que sigui un procés que es pugui desenvolupar independentment fora de l'entorn dels Sistemes d'Informació Geogràfica. (Lagrange, Ruas 1994:1099)

La generalització ha de permetre, en el nou entorn, manipular separatament la "vista" del model de la vista cartogràfica, així com obtenir "vistes" múltiples de les dades geogràfiques a diferents escales i nivells de resolució a partir d'una base de dades idealment única.

La generalització ha de permetre l'accés a les dades a través del que Weibel anomena una "reducció controlada de les dades" en els dominis espacials, temàtics i/o temporals. I ha de permetre la "derivació", a partir de bases de dades, amb graus múltiples de correcció i resolució" (Weibel 1995b:57). L'accés a les dades ha de poder integrar dades que tenen escales diferents per convertir-les a una escala comuna, i és aquesta una de les qüestions centrals en el tractament de les dades, perquè no se sap quina és l'escala de sortida de les bases de dades. La informació, més concretament les dades, planteja un problema d'accés social i tècnic. L'accés per part dels ciutadans a les dades és un espai desigual on no tothom té els mateixos drets; això es veu incrementat per la dificultat tècnica del seu tractament o la varietat, que en molts casos fa que es consideri que la informació està xifrada i, per tant, inutilitzable per el usuari del carrer.

La generalització hauria de permetre una millor comprensió dels fenòmens, a través dels nivells més alts d'abstracció. Actualment, perquè pugui automatitzar-se, la generalització afronta el problema d'haver d'"entendre" a quines escales ocorren els processos espacials; la generalització, en aquest context, pot ajudar a facilitar l'anàlisi: "...the necessity to understand at which scales or range of scales spatial processes occur is one of the driving forces behind generalization today" (Muller 1991:457).

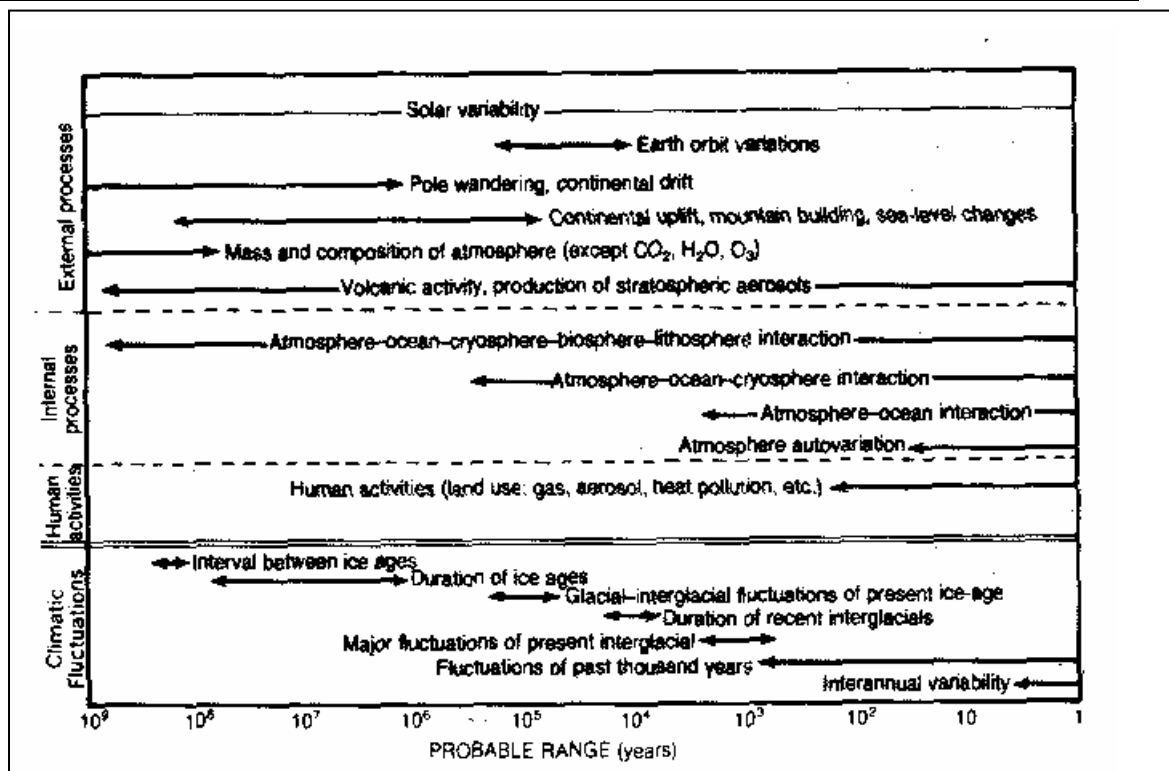
Normalment, els SIG disposen de capacitat d'anàlisi, a través de la connexió amb paquets externs (com, per exemple, un paquet d'anàlisi estadística) o bé a través de la integració d'aquestes funcions. Una gran part d'aquests coneixements procedeixen de

l'escola quantitativa i de la estadística espacial.

Els SIG han anat recollint i acoblant coneixements que no provenen únicament del camp de la geografia. En un principi la intervenció d'especialistes no procedents de la tradició de les ciències espacials ha prioritzat, durant anys, un enfocament basat en una representació de l'espai més mètrica que conceptual, més lligada al camp de la topografia que al de la geografia. La dificultat per resoldre qüestions bàsiques i alhora essencials de la funcionalitat del sistema ha fet també que el sistema d'informació estiguessin dominats per llenguatges anomenats de baix nivell i que hagi calgut temps i esforços per aconseguir que els usuaris dels SIG poguessin interaccionar a través d'interfícies més pròximes als llenguatges naturals de l'usuari, que no pas al llenguatge de màquina.

Els SIG són capaços d'entrar dades, emmagatzemar-les, treure'n sortides gràfiques i fer-ne l'anàlisi per poder-ne extreure informació. Però tant l'anàlisi com la representació cartogràfica requereixen els coneixements existents suplementaris, que no té l'usuari no especialista. Això fa que hi hagi d'haver un altre especialista, no un enginyer sinó un analista espacial i cartògraf experimentat com a intermediari entre l'usuari i les dades, i aquest és el principal repte que tenen els nous sistemes.

El fet que a certes escales certs fenòmens tenen importància i els mateixos fenòmens a altres escales no en tenen és molt clar en l'esquema de Uninayar.



Font: Unninayar 1988:357

3.1. El sentit geogràfic en els objectes digitals espacials

Sense l'èmfasi que es dona a la component espacial, tant per el que fa a la forma espacial com a la "localització" dels fenòmens geogràfics, és difícil d'imaginar el seu tractament tant en els sistemes d'informació geogràfica com en la mateixa cartografia. Quan ens referim a la forma espacial d'un fenomen, ho fem referint-nos a la seva dimensió 0,1,2,3-D i a la seva forma.

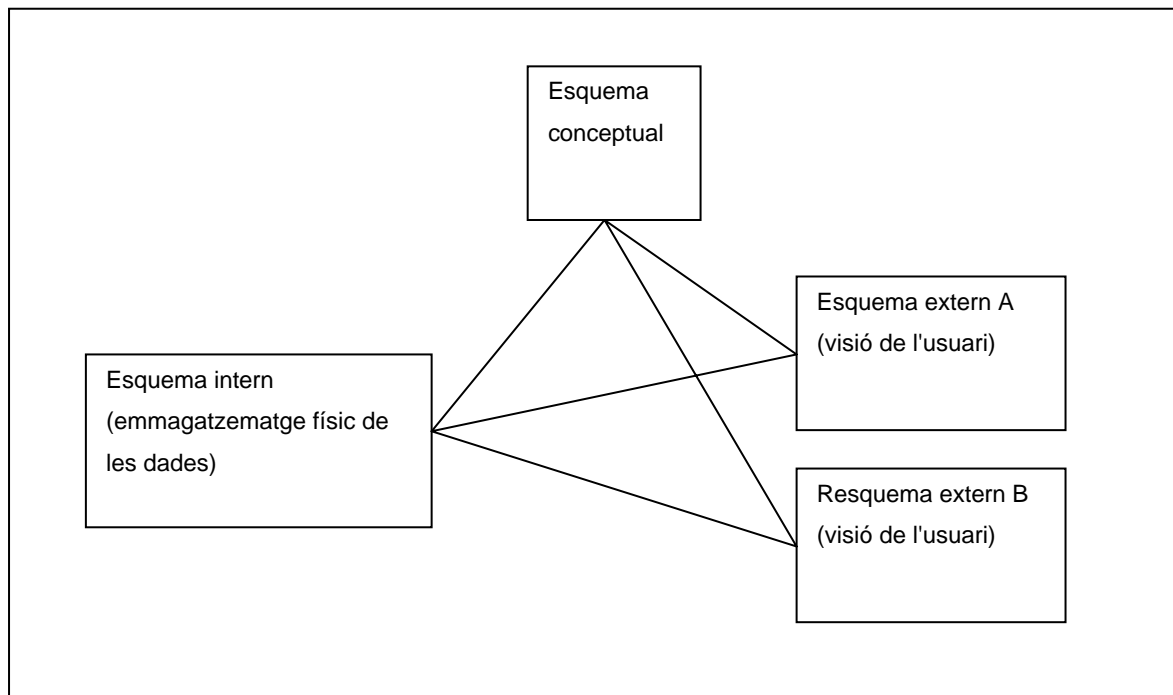
Les dades representen informació del món real en algun moment² i en algun lloc, a través de la referenciació dels objectes a un espai i de la seva caracterització mitjançant uns atributs.

Els fenòmens geogràfics o entitats que triem per derivar-ne coneixement del món real, els fem correspondre a uns elements que anomenem els primitius gràfics o objectes

- _____

² El que centra el debat en aquest camp es com arribar a definir la unitat individual o unitat base (MAUP) i no solsament en termes espacials sino també temporals. ¿com varia una unitat al llarg del temps? No es només la variació espacial la que ens pot interessar sino la variació

espacials. El procés d'abstracció, no sempre considerat part integrant de la generalització i que mereix ser-ho, és d'una gran importància en la fase inicial del procés de compilació cartogràfica. Representa la conversió i codificació de la realitat en la base de dades com una visió determinada del món. De la identificació dels elements que reconeixem i seleccionem del món i de la seva classificació, en depenen la resta d'operacions, tant analítiques com les que tenen a veure amb representacions posteriors. Quan hem de decidir quin tipus d'anàlisi fem, és d'importància primordial conèixer quin ha de ser el procés d'ordenació en la caracterització de les variables geogràfiques en categories o classes. ([Walford Nigel, 1995 #489]:16)



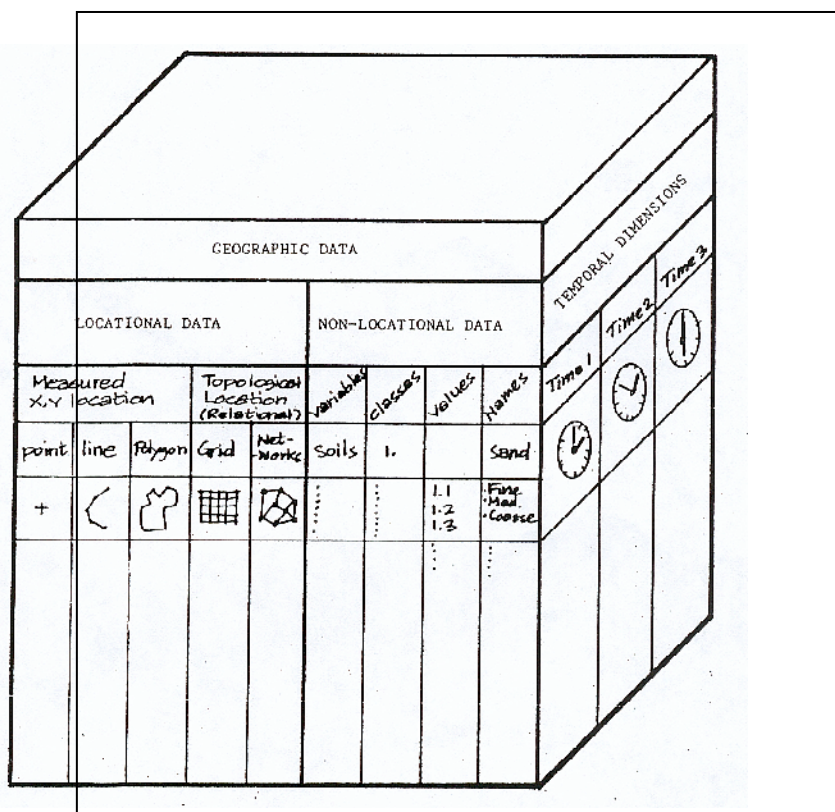
(Font: Frank, Mark 1991:148)

El món és infinitament complex, i el contingut de les bases de dades representa una particular visió del món; l'usuari veu el món real a través de la base de dades. Les mesures i les mostres contingudes en la base de dades han de ser doncs tan complertes i acurades respecte a la realitat com sigui possible. La base de dades ha de contenir els temes i les característiques, el període temporal cobert i l'àrea d'estudi, per tant, és

• _____
en el temps.

important la manera com s'ha fet el mostratge pel que fa al grau de correcció i als estàndards desitjables.

Els fenòmens del món real, els podem observar segons tres dimensions: l'espacial, la temporal i la temàtica, tal com diu Sinton (1978, citat per [Parson, 1990 #481]). Quan observem els fenòmens del món real, normalment el que fem és fixar una dimensió, variar l'altra de manera controlada i mesurar la tercera. Això és important ja que la dimensió en la qual s'emmagatzema la informació en una base de dades influeix decisivament sobre la definició dels problemes i sobre la manera de solucionar-los. Dangermond aporta una variació, afegint en la caracterització de les dades geogràfiques la distinció entre els elements locatius i els no-locatius o atributs, equivalents a la manera de modelitzar l'espai com a vector o com a tessella.



Els tres components d'un Sistema d'Informació Geogràfica segons Dangermond (Font: NCGIA 1990)

Les dades i la informació que en destriem són importants, i així ho mostren Davis i Simonet ([Davis, 1991 #498]:191), que puntualitzen la diferenciació entre els següents termes:

- mesures: recollides amb aparells especials com ara els sensors remots,

informació geogràfica: saber corresponent a un determinat fenomen geogràfic

informació cartogràfica: informació geogràfica obtinguda a partir dels mapes

- dades: mesures i informació que introduïm per a l'anàlisi geogràfica.

3.1.1. L'espai mètric i l'espai geogràfic

Quan parlem de dades espacials és necessari referir-nos a l'espai on es defineixen aquests objectes. Podem definir l'espai, seguint Peuquet, com la relació que s'estableix en un conjunt d'objectes, també anomenats entitats o fenòmens del món real (Gatrell A C, 1991 #486]:119), als quals s'associen atributs o propietats.

El concepte d'espai en un sistema d'informació és difícil de formalitzar, ja que el llenguatge comú no es correspon totalment amb la terminologia, més rígida, utilitzada en els SIG. En el procés de comunicar és necessari poder relacionar els conceptes espacials que utilitzem les persones amb la definició formal d'espai utilitzada en el llenguatge dels ordinadors. El conjunt de conceptes que utilitzem per organitzar la nostra percepció de l'espai ha de poder trobar la seva correspondència en la interfície. Per tant, hem de poder separar els conceptes lligats a l'esquema conceptual dels que corresponen a l'esquema "extern" o vista dels usuaris de l'esquema intern d'emmagatzematge físic de les dades. Aquesta confusió entre conceptes i implantació física de les dades ens fa confondre *el que* (what) fa un sistema i la manera *com* (how) ho fa. I això és important, especialment en l'entorn dels SIG on la innovació és contínua, i per tant, els aspectes cognoscitius i de llenguatge més relacionats amb la interacció entre el nivell conceptual i la "vista de l'usuari" han de poder tractar-se independentment de l'esquema físic intern d'organització i d'estructuració de les dades. El sistema cognoscitiu de les persones conté moltes ambigüitats pel que fa a la concepció de l'espai, cosa que fa difícil de formalitzar-lo en un programari. El concepte d'espai, tal com assenyala Frank (Frank, Mark 1991:148) és difícil de formalitzar, ja que podem concebre l'espai sota un punt de vista kantian, en què els objectes omplen

l'espai i per tant l'espai és un atribut dels objectes. Un altre punt de vista, el cartesià, seria el que correspon a veure l'espai i les seves propietats com a atributs dels objectes. En realitat, utilitzem cada una d'aquestes visions indistintament, i passem de l'una a l'altra segons la nostra conveniència.

A un model conceptual adequat al tipus de fenomen que ens interessa analitzar, li fem correspondre una representació digital o estructura de dades determinada. Però una estructura de dades pot ser adequada a diversos models conceptuais i tenir a la vegada diverses estructures associades. Aquest fet és desconcertant per l'usuari, i Gatrell, fent referència a la naturalesa de l'espai, assenyala que els sistemes d'informació geogràfica porten associada una visió del món que es basa en una representació euclidiana de l'espai, que pot ser útil per a la investigació, però és més difícil d'entendre per l'usuari normal, molt més familiaritzat amb els fets discrets (Gatrell 1991:112). En l'espai euclidià es prioritza l'espai mètric per sobre del topològic, on la contigüïtat i no la localització és l'element base del sistema de referenciació. Precisament, és en funció d'aquesta manera de veure l'espai, que s'han estructurat els SIG *raster* i vectorial, i amb totes les conseqüències conceptuais que en deriven.

L'espai mètric correspondria a la descripció de les relacions entre distàncies, i és el nivell més baix d'abstracció. L'espai topològic utilitza l'existència de relacions espacials entre punts en l'espai, i el nivell més alt correspon a l'espai estructural (conceptes de teoria de grafs) que tracta només d'entitats i relacions procedents del treball de Sowa (1984, citat a Müller et al. 1995b:6).

La discretització en la base de dades la fem tant de l'espai com del temps, ja que, com assenyala Langran, a més de la distància mètrica, hem de considerar la intervenció del temps.

“...cartographic time is newtonian: space and time do not interact [...] current Gis methods [...] should facilitate the upgrade of atemporal systems to temporal ones. (Langran 1992: 73)

3.1.2. La selecció de la informació espacial: el procés d'abstracció dels fenòmens del món real

En la cartografia americana s'ha considerat que la selecció o no dels fenòmens del món

real és un procés a part de la generalització i dependent exclusivament del propòsit del mapa i de l'escala (Robinson et al. 1984, Shea, McMaster 1989).

Com assenyala McMaster, la selecció pot no considerar-se pròpiament com un subprocés de la generalització, ja que no suposa la modificació de l'objecte. El procés de selecció, en la tradició cartogràfica europea, en canvi, sí que s'ha considerat part integrant del procés de generalització. Aquesta manera de veure el procés d'elaboració del mapa i especialment la compilació de les dades inicials, no tindria en aquest context gaire importància si no fos perquè s'ha fet evident que aquest és un dels subprocessos més difícils d'automatitzar. De tota manera aquest és un tema complex en què fins i tot un mateix autor pot mostrar diversitat d'opinions. Per exemple, Armstrong (Armstrong 1991) cita Robinson *et al* (Robinson et al. 1984) per a subratllar-ne la idea que " ...the selection of objects to be mapped is not part of the process of generalization", i en el mateix volum Nyerges (Nyerges 1991) assenyala que "...in cartography the term generalization commonly applies to selection, simplification, classification, induction and symbolization and should be prefaced with the term "map" or "cartographic". En la 5^a edició (Robinson et al. 1984), tanmateix, es decanta per considerar dins la selecció de la generalització.

En paraules de Robinson, els elements de simplificació, classificació, simbolització i inducció ja inclouen la selecció, i es podrien considerar una generalització de la generalització ([Robinson, 1984 #476] :126). Per exemple, dels elements de la generalització considerats en la cartografia tradicional, la simplificació ja conté un procés de selecció del món. El procés de selecció comporta un procés d'eliminació; si considerem que la funció d'eliminació és la que decideix què s'ha de conservar i què s'ha d'eliminar, podem deduir que es fa la representació dels elements seleccionats.

La importància de tractar la selecció com a part integrant del procés de generalització es deu al fet que el procés més subjectiu i que en l'establiment del llenguatge cartogràfic en què l'intent de donar un aire d'objectivitat al tractament de les dades ha estat primordial, s'ha considerat que la selecció era un tema polèmic, i extern al procés d'elaboració del mapa. La generalització digital, basada en criteris de tractament fonamentalment geomètric de les representacions cartogràfiques, en què s'ha pretès que no hi hagués intervenció del cartògraf, per donar un aire d'objectivitat al procés. Però

és evident que la subjectivitat no prové únicament del tipus de tractament. Una de les operacions de simplificació de les formes geomètriques és la "radical law" de Töpfer (1973), que ens permet reduir el nombre d'elements quan es passa d'una escala de compilació a una altra. El principi bàsic de la llei depèn de la relació entre l'escala de compilació i l'original: no té en compte ni la dimensió dels elements ni la de les diferents escales, únicament la relació entre elles. És a dir, que el tractament és el mateix per a escales que mantinguin una proporcionalitat sense tenir en compte la naturalesa de les dades, ja que la llei no pot precisar quina quantitat de detall o d'elements han de conservar-se. Aquest balanç entre el que es selecciona i el que s'exclou fa que la tria última, tot i aplicant la llei de Töpfer, s'hagi de basar en la selecció que fem dels elements i s'hagi de dur a terme, per tant, d'una manera subjectiva.

Richardson i Muller proposen una generalització integrada, més semblant al que fa un cartògraf en el procés de generalització manual, de les activitats de: simplificació, desplaçament, exageració, combinació, selecció, classificació i simbolització; per aquests autors, els processos de selecció i de classificació estan estretament lligats al contingut temàtic dels mapes, i la intervenció del cartògraf depèn del tipus d'informació que s'hagi de presentar.

Per tant, podríem diferenciar les activitats de la generalització, més dirigides al tractament de la forma en el tipus de representació, de les dirigides a intervenir en la generalització del contingut, com ara la selecció i la classificació. Aquestes activitats de generalització estan molt entrelligades, ja que la competició per l'espai hi ha objectes que desapareixen objectes o que canvien de forma per tal de poder aparèixer ([Richardson, 1991 #147]:136).

La selecció també és un dels elements clau, i Nyerges diferencia entre abstracció de la informació geogràfica i generalització cartogràfica. Nyerges suggereix que s'utilitzi el terme *generalitzaci*, quan ens referim a la generalització de la base de dades, en el model de les dades com a procés d'abstracció del món, i el terme *generalització cartogràfica* quan ens referim al context cartogràfic. La generalització en el model de les dades hauria de poder recollir informació geogràfica dels fenòmens seleccionats. Ens és més fàcil d'integrar aquest tipus de saber en la representació cartogràfica, que no pas en les bases de dades. Normalment, tendim a confondre la informació impresa amb

la realitat o a sobrevalorar la primera respecte a la segona. És important, en aquest sentit, de recordar Board (1984) quan diu que "...[the] geographical meaning is drawn from meaning of the referents of the world, rather than from symbols that appear on a map" (citat a Nyerges 1991:60).

Nyerges assenyala que el concepte de *generalització* és tractat en la literatura d'enginyeria de *software* en un sentit més restringit que en la literatura cartogràfica. En la primera es considera que la generalització forma part de l'abstracció de la informació geogràfica i és compon de: classificació, associació, generalització i agregació (Nyerges 1991:66). En aquest context, l'abstracció correspondria en bona part al que, en cartografia, es considera "selecció", és a dir, a una eliminació de fenòmens de la realitat, seguida de tot de transformacions de la informació que porten a una representació o modelització de la realitat.

"It is the task of a cartographer as a database designer to provide enough meaning to convey (a nature of some) reality. In providing more meaning, a database designer most likely would be enhancing the database since more meaning reflects more (of some) reality"

És, per tant, necessari que el tractament de les bases de dades geogràfiques inclogui un saber amb significació geogràfica "*knowledge of geographical meaning*", i això s'ha d'aconseguir amb un sistema de gestió de les dades que permeti integrar aquest tipus de coneixements a partir de les abstraccions de la informació geogràfica.

"...data management capabilities must be enhanced with knowledge representation capabilities [...] this can be accomplished through enhanced data management architectures that make use of geographical information abstractions" (Nyerges 1991:60).

Les abstraccions de la informació geogràfica consisteixen tant en les definicions com en les relacions del contingut conceptual de les bases de dades.

No s'ha dedicat gaire atenció, fins recentment, al contingut conceptual de la informació geogràfica en el sistema d'informació geogràfica, com ho mostra el fet que utilitzem, com a representacions dels fenòmens geogràfics, els elements gràfics primitius de punt, línia i àrea, d'una banda, i els seus corresponents atributs, de l'altra, sense cap mena de

l·ligam conceptual entre aquests. La informació de tipus gràfic i numèric es troben separades en les bases de dades, i la interconnexió entre l'una i l'altra només s'hauria de poder fer a través d'una formalització a un nivell d'abstracció més alt que el d'ara.

"By focusing on conceptual content, as higher order data composed of definitions and relationships, it is possible to develop databases that are "intelligent" (Parsaye et al., 1989 citat per Nyerges 1991:61)

Les dificultats per avançar en la generalització digital serien doncs, precisament, les mateixes que hi ha per construir un saber sobre el significat geogràfic "*knowledge of geographic meaning*". Així s'hauria perdut el saber sobre el procés d'abstracció de recollir el sentit geogràfic i saber-lo integrar en una base de dades geogràfica.

"...knowledge of the abstraction process in developing a geographical database [...] geographical information abstractions as representations of meaning convey what has been missing, lost, or forgotten in many database models" (Nyerges 1991:62)

El contingut conceptual que descriu el significat dels elements gràfics primitius permet desenvolupar unes estratègies; és el que anomenem "metaknowledge", i la manera d'aplicar-lo és el que s'anomena "metaplanning" (Armstrong 1991:88). El cartògraf utilitza també un tipus d'estratègia basada en el coneixement quan generalitza, per exemple, quan fa tasques difícils de formalitzar, com ara el pas d'un nivell de percepció a un altre. Això ho fa, ja que pot fer ús de la descripció inicial com també del seus coneixements de les convencions cartogràfiques i de les interaccions amb la seva visió del món real.

"It was possible for a cartographer to vary the criteria for interpretation and thus the definition from feature to feature, depending upon the context." [Morrisón, 1994 #328]:6

El cartògraf fa servir tant la informació de què disposa com el coneixement de les regles per generalitzar, i és aquest coneixement el que li permet saber en quin moment, ha de parar el procés de generalització, ja que ha aconseguit el seu objectiu. (Lagrange, Ruas 1994:1100).

3.1.3 Les fonts de dades espacials: analògiques i digitals

Les dades primàries són les dades espacials que es mesuren directament en el camp, i

és la densitat de mostratge la que determina la resolució d'aquestes dades. Es fa servir una mostra per establir la variació present tant a nivell espacial com temporal.

Les dades de les fonts secundàries procedeixen de mapes, taules o altres bases de dades. És important destacar que cal aconseguir informació addicional de les dades mateixes, tant del procediment que s'ha utilitzat per recollir-les com dels processos de compilació que han seguit les dades, així com dels esquemes de codificació del grau de correcció i la precisió dels instruments.

En l'entrada de dades a l'ordinador seria molt còmode que poguéssim conservar les estructures fenomenològiques a què l'usuari està més acostumat (Burrough 1986:15), però generalment això no és possible, ja que probablement no coincideixi la millor manera de visualitzar les dades amb la millor manera de estructurar-les en la base de dades.

Una base de dades consisteix en la representació digital d'objectes discrets que corresponen als fenòmens que després representem cartogràficament. Molts dels fenòmens que apareixen en un mapa són ficticis i no existeixen en el món real, i per tant una base de dades espacial inclou les versions digitals d'objectes reals, les versions digitals de fenòmens cartogràfics artificials com ara les corbes de nivell. I a més, conté objectes artificials creats especialment per a la base de dades com ara els píxels (NCGIA 1990).

Decidint si donem un tractament continu o discret a la informació espacial, determinem les possibilitats posteriors de codificar-la. A continuació, la informació espacial pot ser convertida en propietat perceptiva qualitativa, o bé ordenada o bé quantitativa. Les variables visuals, basades en aquestes propietats perceptives ens han de permetre posteriorment la tria de l'objecte digital cartogràfic.

Tipus d'informació	propietat perceptual
qualitativa	associativa (+/- selectiva)
ordenada	ordenada
quantitativa	quantitativa

En primer lloc, hem de considerar si la informació espacial és present o no, o sigui, que considerem la dicotomia presència i absència d'un element en una àrea determinada com una mesura d'escala. La informació qualitativa correspon a les dades de tipus categòric, que representen una mesura escalar, on un element pot ser classificat en diferents categories. La informació qualitativa permet comptar el nombre d'elements, o el nombre de vegades que un fet ha aparegut en un lloc. En la informació ordenada el rang és representat per la manera d'ordenar els elements en funció d'un determinat criteri.

3.1.4 Els objectes espacials

En un context digital, els mapes són les representacions d'una realitat geogràfica que ha estat enregistrada en una base de dades. Si representem aquesta base de dades geogràfica mitjançant símbols, obtenim una base de dades cartogràfica (Nyerges 1991:60).

Intervenien dues menes d'elements en la base de dades que modelitzem en la base de dades: entitats i objectes espacials. L'entitat és l'element de la realitat, un fenomen d'interès o un procés de la realitat que no és subdivisible en altres fenòmens del mateix tipus com, per exemple una ciutat no ho és en altres ciutats. L'altra mena d'element és la manera de representar aquesta entitat en la base de dades com a objecte. Un objecte és una representació digital de tota entitat o d'una part, i el mètode de representació digital d'una entitat o d'un fenomen varia en funció de l'escala i de la finalitat principalment. Això és el que s'anomena "representacions múltiples": la possibilitat que posseeix un fenomen únic en el món real de poder ser abstret, interpretat i representat de diferent manera, a diferents escales, mitjançant diferents elements gràfics. Si aquest objecte, el representem cartogràficament com a element del mapa, és un objecte digital cartogràfic.

La naturalesa de la informació ens ha de permetre que en el pas de la realitat al model cartogràfic de les dades hi hagi la màxima correspondència morfològica entre la percepció que es té d'un fenomen i la codificació que se'n fa a la base de dades.

"...[an] approach to a typology-syntactics for quantitative spatial data is there should be an iconicity in spatial structure between a reflect phenomena and the map" (MacEachren 1995: 303)

Les dades, en el model de les dades, que anomenem objectes digitals, no tenen per què coincidir amb l'element gràfic primitiu cartogràfic que utilitzem posteriorment en el model cartogràfic: en el model de les dades les necessitats de l'anàlisi tenen prioritats i qualsevol tipus de codificació no permet qualsevol tipus d'anàlisi. Així mateix, és molt possible que no mantinguem el mateix tipus d'objecte digital, almenys en la seva dimensionalitat, en el model de les dades i en el cartogràfic. Per exemple: a un fenomen o entitat, podem donar-li un tractament de punt en el model de les dades i representar-lo posteriorment cartogràficament a través d'una àrea. Les regles de codificació en el model de les dades i en el model cartogràfic es basen en tècniques diferents, en l'aplicació de l'estadística espacial en el primer, i en el tractament de les dades a través de les seves representacions cartogràfiques en el segon. La possibilitat d'accedir a una tècnica molt potent capaç de tractar grans volums de dades, no ens ha de fer oblidar que el tractament de les dades, i de la informació, a través de la seva representació cartogràfica ha estat una manera molt intel·ligent, sense ordinadors, de manipular simbòlicament els fenòmens corresponents al món real. En els SIG hi ha hagut errors de codificació i tractament de les dades, especialment en el pas de la codificació del món real a la representació de les dades en el model cartogràfic; per desconeixement de l'usuari de les regles cartogràfiques de manipulació i per l'aparent facilitat per realitzar aquestes transformacions, sense que els programes s'hi oposin. Aquests errors han portat a conclusions errònies de la informació resultant. Amb això, intento recalcar que el llenguatge cartogràfic, que es va basar en la *Gestalt* i que s'intenta replantejar amb els nous avenços del coneixement perceptiu, pot doncs funcionar com a sistema de referència, és a dir, se'l pot seguir o criticar, però no se'l pot desconèixer com fa l'usuari d'una banda, i els venedors de SIG, de l'altra.

Els Sistemes d'Informació Geogràfica ens permeten transformar els objectes espacials o crear-ne de nous per augmentar-ne les possibilitats analítiques, com per exemple la creació de superfícies a partir de punts, i per tant és obvi que no s'ha de perdre el control del procés, en especial, com es diu més amunt, quan s'adjudiquen codificacions i es donen tractament a dades que per la seva naturalesa no poden ser objecte de qualsevol tipus de tractament que aparegui en els programes de cartografia o dels SIG.

3.1.4.1 Tipologia dels objectes espacials

El National Standard for Digital Cartographic Databases classifica els objectes espacials segons la seva dimensionalitat espacial.

Els elements gràfics primitius als quals fem correspondre els objectes digitals són: el punt, 0-D (sense dimensió); la línia de 1-D, que fem correspondre a fenòmens lineals; l'àrea de 2-D, per representar àrees; i la superfície de 3-D, que permet, mitjançant la interpolació de punts, crear superfícies estadístiques, un tipus de tractament de les dades espacials que en la cartografia tradicional era més difícil d'utilitzar per la necessitat de computació.

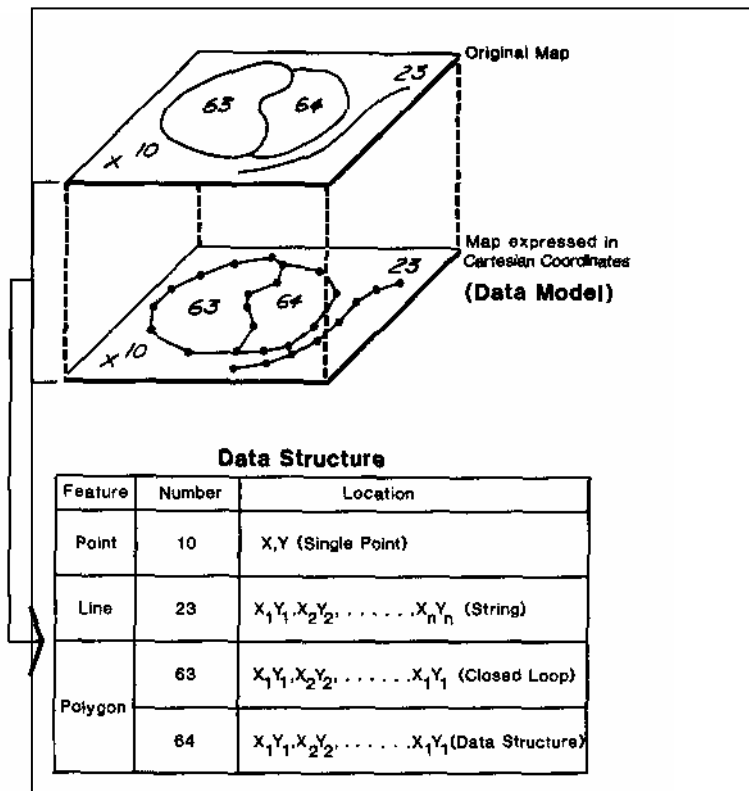
Definim normalment els punts, línies i polígons en els mapes utilitzant les coordenades cartesianes x, y com en el cas de la longitud/latitud de l'espai euclidià. El càlcul a través de les coordenades cartesianes és el mètode més comú per mesurar la localització espacial i analitzar les propietats de la localització en l'espai. Un altre mètode consistiria a representar l'espai mitjançant la teoria de grafs, que utilitza la relació topològica entre els elements o les localitzacions relatives. Una altra tècnica per representar les dades geogràfiques és la utilització de cel·les. Aquests dos mètodes d'identificació espacial de la informació corresponen als dos models: el vectorial, o de representació del fenomen a través d'una polilínia, i el *raster* o tessel·la o de cel·la.

Aquests dos models són limitadors per a la generalització i en especial per als operadors que intervenen en les tasques de generalització³, a causa, en bona part, de la tipologia dels objectes espacials en funció de punts, línies i àrees, que són un model gràfic primitiu (Tobler 1988).

En l'espai de coordenades cartesianes definim els elements geomètrics de punt, línia i polígon, que permet basar-nos en la geometria euclidoana i per la inclusió d'aquests en la base de dades com a objecte digital, com mostra la figura següent:

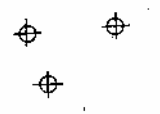
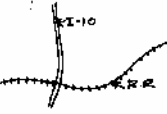
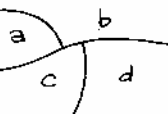
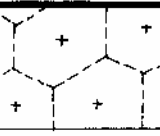
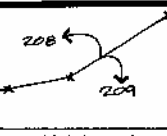
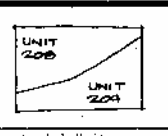
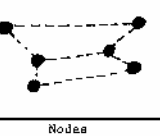
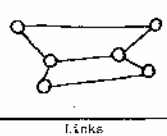
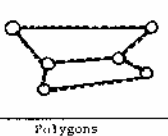
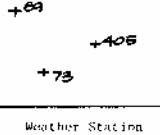
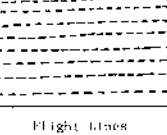
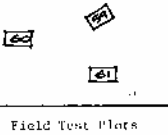
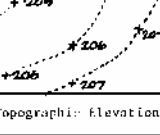
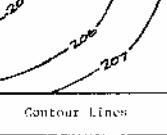
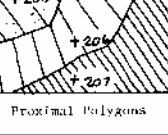
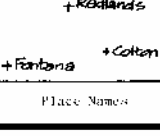
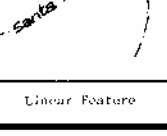
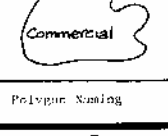
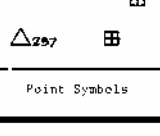
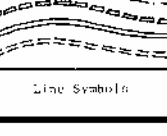
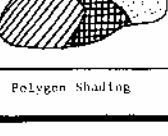
- _____

³ De totes maneres, s'ha fet servir en anàlisi espacial durant anys (Unwin 1981), i a partir de les definicions de Robinsón ([Robinson, 1984 #476]) i Hagget (1965).



Font: Dangermond 1988:344

Si considerem no només el mètode d'identificació espacial del sistema de coordenades cartesianes sinó també la possibilitat de considerar l'espai topològic, podem afegir als elements geomètrics de mesura de punt, línia i polígon, les seves relacions topològiques tal com es representa en la figura següent.

	POINTS	LINES	POLYGONS
FEATURE DATA	 Point Feature (Archaeological Site)	 Linear Features (Roads)	 Homogeneous Polygons (Soils)
AERIAL UNITS	 Polygon Centroids	 Administrative Polygon Boundaries	 Aerial Unit (Census Tract)
NETWORK TOPOLOGY	 Nodes (Intersections)	 Links (Streets)	 Polygons (Blocks)
SAMPLING RECORDS	 Weather Station	 Flight Lines	 Field Test Plots
SURFACE DATA	 Topographic Elevations	 Contour Lines	 Proximal Polygons
TABLE/TEXT DATA	 Place Names	 Linear Feature	 Polygon Labeling
GRAPHIC SYMBOL DATA	 Point Symbols	 Line Symbols	 Polygon Shading

Font: Dangermond 1988:347

En l'actualitat es plantegen models de dades més apropiats com ara les representacions espectrals, els models de Voronoi o bé els models de dades jeràrquics que s'adapten bé al canvi d'escala i al sistema classificatori predominant en les ciències de la terra (Weibel 1995a:260)

Els sistemes orientats als objectes permeten definir de manera diferent al concepte d'objecte espacial i d'agrupar els objectes en "classes", "tipus" i "objectes complexos"

(s'on els que construeixen d'altres objectes) com també "l'encapsulació" de la propietat "d'herència" de les dades i dels processos, i la transmissió dels canvis. Els anomenats "objectes complexos" permeten la construcció d'un model de les dades en un entorn orientat a objectes, més pròxim al món real, encara que tenen el greu inconvenient de manca d'estandardització en les definicions dels conceptes i terminologies.

3.1.4.2 Tipologia de les relacions espacials

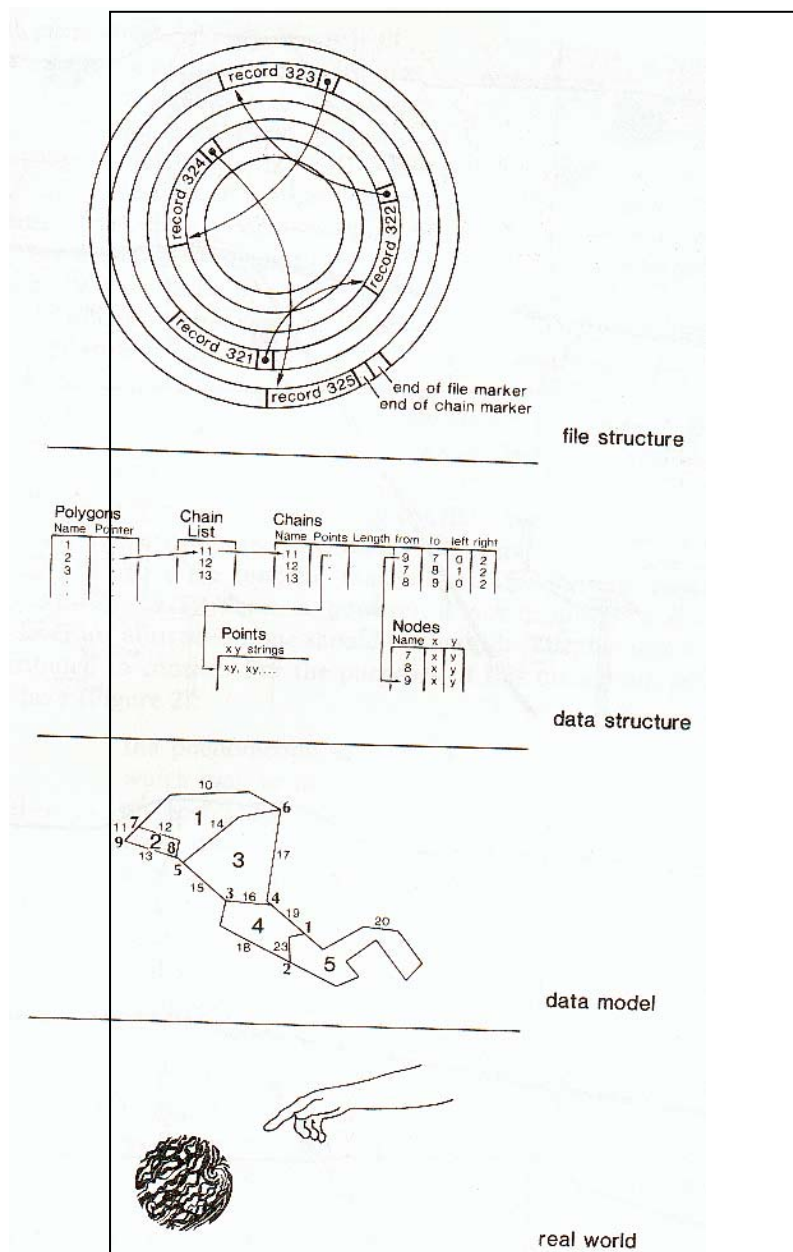
Les relacions de les dades espacials poden ser de tres tipus: relacions que s'utilitzen per construir objectes complexos a partir dels elements gràfics primitius, relacions que es poden calcular a partir de les coordenades dels objectes i relacions que no es poden calcular des de les coordenades i s'han de fer durant l'entrada.

"It is the user of a relational database who holds the semantic information, not the relational database" (Oxborrow 1989:270).

En el l'enfocament fenomenològic de Mark (Mark 1989:76) s'insisteix en això: que la generalització ha de tenir en compte la naturalesa del fenomen en el món real que representem simbòlicament a través d'un objecte gràfic. Aquesta idea no és realment nova, ja que el 1965 Imhof sostenia això mateix, i tanmateix aquesta és una idea que s'ha perdut.

3.1.4.3 Estructuració i models de les dades espacials geogràfiques

Frank i Mark assenyalen el fet que per els científics de la computació s'utilitza el terme de "model de dades" com eines i mètodes a la vegada, que permeten descriure la realitat o la nostra percepció de la realitat ho fan diferentment en alguns casos del com es fa en la literatura dels SIG, a on *Raster* i *Vector* corresponen a la estructura de les dades en un model espacial determinat (Frank, Mark 1991).



(Font: Peuquet 1990:250)

Les dades espacials estan geo-referenciades, representades i emmagatzemades en un entorn digital; segons Frank i Barrera (1990) hi ha quatre maneres de referenciar aquestes estructures espacials entre si:

- 1) segons el tipus geomètric de dada (de punt a regió)
- 2) segons el tipus d'objecte (de no fragmentades a fragmentades)
- 3) segons l'accés a les dades (de directe a jeràrquic)
- 4) segons la subdivisió de l'espai (de regular a determinat per les dades)

Normalment, en un sistema d'informació geogràfica, les dades estan representades a través d'un sistema de cel·la *raster* (estructura de regió fragmentada, accés directe, determinat per les dades) o de vector (regió no fragmentada, accés directe, determinat per les dades). Hi ha incompatibilitat principalment entre els sistemes que utilitzen dades de tipus cel·la procedents dels sensors remots, i els que n'utilitzen amb format de vector, procedents de la cartografia.

3.2. Entendre el preprocessament

A l'article que ha estat el marc de referència del procés d'automatització en la generalització (Brassel, Weibel 1988:231), Brassel i Weibel defineixen el procés de modelització de les dades en un SIG, i estableixen que per poder generalitzar, necessitem basar el processament la comprensió o intel·lecció del procés. Recentment, d'altres autors coincideixen en aquest punt: la necessitat de comprendre i dominar el que fem quan utilitzem els SIG, moltes vegades difícil de controlar en "la caixa negra" de l'ordinador. Així trobem que la "precartografia" de Muller o el procés "d'abstracció de la informació geogràfica" de Nyerges o el "saber geomètric i estructural" d'Armstrong són equiparables: aquests processos s'han de desenvolupar durant el preprocessament dins del model de les dades.

La comprensió del procés, pròpia d'una persona amb experiència, és difícil d'automatitzar; això, de moment només s'ha aconseguit aplicant una successió mecànica de funcions, o bé d'algoritmes de manera aïllada. La comprensió és el que resultat de "donar un cop d'ull" o fer una inspecció visual de la informació original, i a partir d'això, de triar un tipus de processament de les dades o un altre. Si volem simular el procés de manera automàtica, hem d'entendre el procés de reconeixement dels fenòmens essencials o importants que ens permeten modelitzar-lo.

"To understand generalization means to extract the essential structures of the spatial information available (in the thematic, spatial and temporal domains), to identify the essential processes for modifying these structures and to formalize these processes of modification adequately as a number of

operational steps" (Brassel, Weibel 1988)

Dins del marc de treball conceptual per a la generalització, el "procés d'entendre" comporta que es reconegui l'estructura, que els operadors reconeguin quines són les funcions necessàries i els criteris, i que es posi en marxa el procés de modelització. L'execució del procés i la presentació de les dades completen aquest marc de treball. El model de Brassel i Weibel té l'avantatge de poder integrar altres conceptes així com permetre una igualtat de processos amb els definits per Armstrong: reconeixement de l'estructura o saber estructural, procés de reconeixement o saber processual.⁴ En el seu treball sobre la classificació dels tipus de coneixement necessaris per generalitzar, Armstrong distingeix tres tipus de "saber": geomètric, estructural i processual; els dos primers corresponen amb el reconeixement de l'estructura de les dades, i el tercer amb el reconeixement del procés, en l'esquema de Brassel i Weibel (Armstrong 1991:89).

El saber que utilitzem per generalitzar, l'extraïem principalment de tres fonts diferents: dels llibres de text i les agències cartogràfiques, de les sèries cartogràfiques, i en tercer lloc, de l'experiència del cartògraf. La formalització d'aquests tres tipus de saber permetria integrar-los en un sistema intel·ligent.

Nyerges critica Brassel i Weibel pel que fa al marc de treball conceptual, ja que aquests consideren que, si es fa sota control estadístic, l'error és minimitzable en la generalització. Nyerges manté que això no és cert, si tenim en compte les diferències conceptuals entre la base de dades geogràfiques original i la base de dades resultant. Nyerges mostra que encara que es facin servir tècniques estadístiques, aquestes no són suficients per controlar la pèrdua de "sentit," i suggereix utilitzar les abstraccions de la informació geogràfica⁵ per controlar-ne la possible pèrdua conceptual durant el procés de generalització estadística, a través del seguiment de la representació de les distribucions espacials amb un llinatge conceptual (*lineage*). Les abstraccions de la informació permeten incorporar "sentit" al procés de derivació d'un DLM en una altre

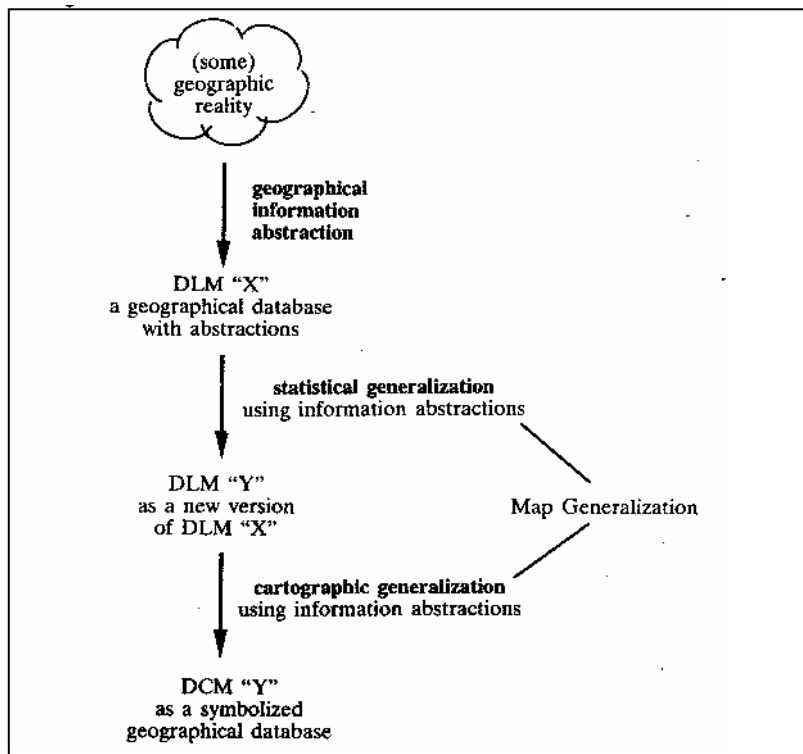
• _____

⁴ Això és de gran interès ja que permet la utilització d'operadors ja provats per altres sistemes (Reichenbacher 1996:13).

⁵ La generalització estadística i l'abstracció d'informació són diferents, sobretot pel que fa al resultat; mentre que l'abstracció de la informació geogràfica està representada en la base de dades i permet la supressió temporal de detall espacial, temàtic o temporal. Els processos de generalització estadística i cartogràfica fan una transformació (Nyerges 1991:64).

DLM, i el llinatge permetria fer-ho en el procés de generalització estadística o cartogràfica.

"Geographical information mainly concerns managing geographical meaning in databases, and map generalization mainly concerns structuring map presentations [...] In fact geographical information abstractions can assist with performing statistical generalization by helping to preserve the meaning of the resultant spatial distributions. This would involve representing concept lineage for spatial distributions. In doing so, selected characteristics are statistically processed, but whether the conceptual meaning of the spatial distributions changes depends to a significant degree on just what characteristics are."(Nyerges, 1991 #144]:63)



(Font: Nyerges 1991:64)

3.2.1. Reconeixement de l'estructura: el saber geomètric i estructural

Una de les raons per les quals els cartògrafs amb experiència no han explicat mai com

es generalitza es deu al fet que no han tingut mai l'oportunitat d'anar especificant de manera sistemàtica el tipus de saber que apliquen a cada diferent tipus de tasca de generalització. La base de dades geogràfica és una representació digital de les concepcions humanes sobre el món i fa referència a aspectes tant espacials i temàtics com temporals que permetin representar un coneixement.

Però en el camí de l'automatització, per intentar formalitzar el saber d'un cartògraf, mitjançant un procés tan clar com sigui possible i "intel·ligent", el principal entrebanc és que actualment les bases de dades no contenen gaire més informació, que la geomètrica, i estan mancades del contingut geogràfic que en faciliti el processament. Cal poder integrar, en el tractament de les bases de dades geogràfiques, un saber que tingui "sentit" geogràfic, i aquest, només el podem aconseguir directament dels referents en el món, no pas dels símbols que apareixen en un mapa.

"It is the task of the cartographer to provide enough meaning to convey (a nature of some) reality. In providing more meaning, a database designer most likely would be enhancing the database since more meaning reflects more of (some) reality". (Board, 1984, citat per Nyerges 1991:60)

Nyerges diferencia dos "sabers", un d'intern compost principalment de les condicions necessàries i suficients dels aspectes espacials, temàtics i temporals de la informació: i un altre d'extern que inclou els aspectes significatius de les relacions entre els fenòmens geogràfics. Aquestes relacions i entitats representades en la base de dades corresponen a les dels fenòmens en el món real. Per Nyerges, la dificultat per desenvolupar la generalització consisteix en això: que el "saber" sobre el procés d'abstracció per desenvolupar una base de dades geogràfica s'ha perdut, o no està actualment a la nostra disposició en els sistemes. Actualment, el saber amb sentit geogràfic, com a part del procés d'abstracció, no és fixat i es desenvolupa la base de dades, només a partir dels elements gràfics primitius: punts, línies i àrees i els seus atributs.

S'accepta en el moment actual que és necessari poder diferenciar entre l'abstracció de la informació i la generalització en l'activitat cartogràfica. Així, en el model de Brassel i Weibel (1988), com en el Nyerges (1991), l'abstracció de la informació consisteix a elimina-hir els aspectes més específics i deixar-hi els més importants de la realitat; aquesta és una típica activitat humana, que ens permet aïllar els aspectes importants del món i representar-los en el disseny de la base de dades geogràfica.

"A representation of meaning for and in a geographical database could amplify the intelligence of data in support of digital map generalization" (Weibel 1991:172)

Les abstraccions corresponents a la informació geogràfica són construccions basades tant en les definicions com en les relacions geogràfiques per poder desenvolupar el contingut conceptual de les bases de dades. El contingut conceptual de les bases de dades és, o hauria de ser, en les metadades, que contenen dades sobre les dades, a la vegada que descriuen el significat dels elements gràfics primitius de les dades. És a partir del contingut conceptual d'alt nivell ("higher-order data"), compost de definicions i de les seves relacions, que ha de ser possible de desenvolupar bases de dades "intel·ligents".

Lagrange i Ruas també estan d'acord amb Nyerges en el fet que el cartògraf, quan generalitza dades cartogràfiques, el que pretén és construir una representació gràfica que correspongui a un nivell de percepció partint d'un altre nivell de percepció.

Quan realitza aquest procés de generalització, utilitza la descripció inicial del món real que coneix, com també la els seus coneixements sobre les convencions cartogràfiques. Això mostra que la informació inicial, necessària per poder generalitzar, no es pot basar només en informacions geomètriques o elements geogràfics molt simples, sinó que necessitem que les dades geogràfiques continguin saber geogràfic, utilitzable en el seu processament. En aquesta línia es trobaria també la necessitat de poder comptar amb objectes que a un nivell superior, en el pla semàntic, permetin generar representacions múltiples, a diferents nivells descriptius, que podem utilitzar simultàniament. Això es d'interès amb vista al canvi d'escala, ja que només s'aconsegueix deduir quines són les representacions múltiples associades a un objecte, si podem definir l'objecte a alt nivell.

La comprensió del procés consisteix a preservar al màxim possible les propietats geomètriques, espacials i semàntiques, tot respectant les limitacions gràfiques en què es basa la simbolització i la resolució espacial (Lagrange, Ruas 1994). El reconeixement de l'estructura permet ajudar a identificar les propietats i les relacions entre objectes digitals procedents del món cartogràfic, com el reconeixement de patrons espectrals que caracteritzen les imatges d'estructura cel·la, procedents de sensors remots. En aquest camp que pertany a la teledetecció el reconeixement dels objectes digitals amb

estructura de cel·la es basa en la discriminació de les signatures espectrals.

El reconeixement de l'estructura de la base de dades original pretén:

- la identificació d'objectes i agregats
- la determinació de les seves relacions espacials
- l'establiment de mesures d'importància

Nyerges segueix l'esquema de Brassel i Weibel, però hi introdueix el concepte d'abstracció de la informació geogràfica, que defineix com:

"Geographical information abstraction constructs are the repository for the conceptual information identified as part of structure recognition." (Nyerges 1991:65)

Les abstraccions representen el saber corresponent a l'estructura de les relacions de veïnatge geogràfic que ens interessa en el reconeixement de l'estructura; i en aquest context, tant un conjunt d'entitats com les relacions associades en aquestes, poden variar segons el context.

Nyerges cita treballs procedents de la percepció i el reconeixement visual que han demostrat que el reconeixement de l'estructura precedeix com a procés cognoscitiu la resta de funcions del procés de compilació de les dades, i assenyala " ...meanings are not generated spontaneously, but evolve from what is already known...". (Nyerges 1991)

La identificació d'objectes segons sigui el propòsit del mapa és un procés de selecció que permet la classificació conformement al comportament estàndard del saber d'un cartògraf. En l'aplicació d'aquest tractament de les dades, considerem els factors corresponents a: la densitat dels fenòmens, la competència d'aquests per l'espai, també anomenada *ratio* blanc/negre (Black/White ratio); la importància del fenòmen; el context geogràfic; el propòsit del mapa i l'escala de reducció. Podem dir que les persones tenim generalment una idea "a priori" de les relacions espacials i la dependència que es dona entre els objectes en l'espai i, per tant, sobre quines han d'ésser les prioritats.

Si considerem que el procés de selecció és un procés subjectiu, dos cartògrafs no farien una mateixa selecció d'elements de la realitat. A través de la possibilitat d'aprenentatge de què disposen els sistemes neuronals, Muller, basant-se en el treball de Richardson, descriu quins són els tipus de fenòmens que apareixen a diferents escales en els mapes

base canadencs. I a partir d'això descriu el procés que permet definir el tipus de saber per poder instruir la màquina que ens ha de permetre fer la selecció: "we have to ask what are the sources of knowledge which can be used to instruct the machine" ([Muller, 1992 #495]:133)

Els mecanismes que regeixen l'abstracció geomètrica, necessaris per crear saber geomètric d'alt nivell, han de diferenciar-se dels mecanismes que regeixen la representació geomètrica cartogràfica. Com que els mecanismes d'abstracció cartogràfica estan ben definits, Muller proposa que s'integrin en el marc més general de la generalització cartogràfica, ja que el treball de Brassel i Weibel és ambigu en la formalització, i aquests autors es basen en un imprecís conjunt d'operacions de generalització gràfiques.

"The approach of abstraction mechanisms might be a major step forward in this field offering well defined and orthogonal "operations" while disregarding the problems of what knowledge about the world to keep and how to communicate it graphically to the humans"(Muller 1992:123)

El mecanisme d'abstracció geomètrica que Muller proposa com a exemple permet ser utilitzat en el disseny d'estructures de dades per abstraccions múltiples dins dels GIS. El món és de tipus continu i infinit, i quan hem de cartografiar-lo, seguint Goodchild el que fem és donar un tractament continu a una "discretització" d'aquest món continu. L'operació de passar d'un continu a un model detallat "base abstraction" és encara infinita, però com que els ordinadors són màquines finites, els mecanismes d'abstracció geomètrica han de representar-se amb models finits, en què el tractament discontinu afecta tant l'espai com el temps. De fet, com podria dir Goodchild, no es pot suposar que la proximitat locativa implica la topològica.

La identificació de l'estructura comporta identificar les propietats i relacions entre els objectes que es necessiten per a la generalització. Aquestes propietats i relacions inclouen característiques geomètriques, topològiques així com relacions espacials i propietats semàntiques.([Ruas, 1995 #488]:76)

El reconeixement de l'estructura està controlat per :

- els objectius de la generalització

- la qualitat de la base de dades original
- l'escala final del mapa
- les regles de comunicació pel que fa als límits tant gràfics com perceptius

El reconeixement de l'estructura està poc desenvolupat i Nyerges proposa que siguin les:

"Geographical information abstraction is the repository for the conceptual information identified as part of the structure recognition. Such abstractions represent the structural knowledge..." (Nyerges 1991:65)

La possibilitat d'identificar de manera automàtica l'estructura de les dades geogràfiques i cartogràfiques és una qüestió crucial per al desenvolupament de la generalització cartogràfica. Buttenfield, en aquest sentit, mira si hi ha dependència entre l'estructura geomètrica de les dades i l'escala. Utilitza com a exemple un element corrent en l'anàlisi de les dades en percepció remota, que és la signatura espectral; a diferents fenòmens, els corresponen diferents signatures estructurals (*structure signature*) (Buttenfield 1991:150), si considerem que el significat d'un fenomen és important en la interpretació de la seva geometria, hem d'admetre que la signatura espectral ens orienta al nivell més abstracte sobre el comportament típic d'un fenomen determinat, cosa que permet deduir quin ha de ser el seu processament geomètric, temàtic o temporal.

3.2.1.1 Necessitats de les estructures i dels models de les dades espacials per a la generalització

És evident que els models de dades cartogràfics han de ser millorats i han de contenir més informació descriptiva que faci referència a la naturalesa de les dades i a l'estructura dels fenòmens representats, ja que això simplifica en gran manera el processament posterior.

El grup de generalització de l'ICA, en la preparació del manual de generalització "GIS and Generalization" (Müller et al. 1995a), a Compiègne, va proposar als participants uns temes per preparar el seminari. Aquests temes es centraven en la recerca per millorar el reconeixement de l'estructura de les dades o "saber estructural", amb

mètodes que permetin la definició i l'extracció de saber estructural. Això hauria de comportar un conjunt de millores en les funcions que permeten recollir el grau de complexitat, la distribució i les relacions espacials dels objectes cartogràfics (*cartographic features*) i també en els operadors que permeten seleccionar i controlar la generalització. Aquest grup de l'ICA avança una hipòtesi de treball: les estructures de les dades i els models de les dades que s'utilitzen normalment són arcaics i no permeten qualsevol tipus d'enfocament: les operacions de generalització depenen doncs del context. Aquestes estructures actuals, com per exemple la corresponent al model de dades topològic, d'estructura d'arc-node, no seran mai capaces de resoldre aquest tipus de conflicte, ni tan sols capaces de detectar-lo .

Els models de dades han de permetre que ens acollim a un marc de referència en què puguem expressar les relacions espacials de manera flexible i clara. Així, a més a més del vectorial i del *raster*, s'haurien de considerar d'altres models com ara el que es defineix en l'espai estructurat en grafs. Els models de dades espacials en un entorn orientat a objectes són una promesa, ja que, com assenyala Muller (Müller et al. 1995b:5), permeten "l'herència" de certs processos i atributs que poden ser utilitzats per a les actualitzacions posteriors que fem de les dades. Però aquesta facilitat no ens soluciona el problema de les necessitats en el moment d'estructurar el procés de generalització a causa del comportament distint dels fenòmens geogràfics a diferents escales.

Així ho assenyala Buttenfield (Buttenfield 1995:91): és feina del cartògraf que compila la informació preservar l'associació del "sentit" geogràfic i cartogràfic; i cita el treball de Richardson (1961), on es fa manifest que les característiques mètriques són dependents de l'escala, i canvien amb la resolució, així com les característiques no mètriques, com ara la toponímia, que no té sentit a totes les escales.

"The user must be able to access information at different levels of detail, from local to global scales, and at various levels of abstraction, from individual objects to the "envelope" of their classes" (Muller 1991:462).

3.2.2. Reconeixement del procés (saber processual)

Una vegada hem reconegut l'estructura de la base de dades original, podem definir els

processos de generalització, o sigui, quins són els operadors i corresponents algoritmes que hem d'aplicar. Aquest pas consisteix en la identificació de diferents processos essencials i en la presa de decisions per modificar aquestes estructures basant-nos en l'establiment de:

- els paràmetres de control
- els tipus de modificacions de les dades
- i els paràmetres de les estructures finals.

Per tant, hem de definir: què fem la base de dades original; quins tipus de conflictes haurem d'identificar i resoldre; quins tipus d'objectes i estructures s'han de realitzar en la base de dades. Això es difícil d'aconseguir, ja que la majoria de SIG estructuren la informació espacial en capes d'informació, i la generalització ha de poder conservar la relació d'un fenomen amb la resta de fenòmens, amb els quals hi ha alguna relació establerta; independentment de que es trobin o no a la mateixa capa d'informació. El reconeixement del procés ens permet adaptar les nostres decisions a un tipus de processament, i en aquesta fase s'ha de tenir doncs tenir en compte quina és la finalitat desitjada, si volem iniciar el tractament en el model de les dades o volem utilitzar el nostre sistema per representar la informació com a model cartogràfic.

Està molt ben establert en la cartografia tradicional el procés que s'ha de seguir en el model cartogràfic de compilació i decisió pel que fa al millor tipus de codificació de les dades. Una de les escoles que té un reconeixement dels processos molt ben establert és l'escola suïssa. I a propòsit del processament de la cartografia tradicional, moltes vegades considerat altament subjectiu, no hem d'oblidar que una crítica injusta és no prendre prou seriosament el preprocessament de les dades. Això té a veure amb el fet que el producte més conegut és el mapa topogràfic, i que tenint aquest molt ben definida la seva finalitat, no ha exigint cap reflexió sobre el tipus de representació sense replantejar-se prèviament quin era el millor tipus de codificació de les dades.

3.2.3. Modelització del procés

Un cop formalitzat el procés de reconeixement de l'estructura, apliquem aquest seguit d'operacions:

- compilació de regles i de procediments que es troben en una "llibreria de processos"

- tria dels paràmetres que s'estableixen en el procés de reconeixement.

Apliquem a la base de dades original les estructures de la informació reconegudes, per després, a través del procés d'execució, les convertim en la base de dades final generalitzada. El procés d'execució seqüencia una sèrie de d'operadors, que es troben en la llibreria, per fer la representació de les dades i convertir-les en el mapa final.

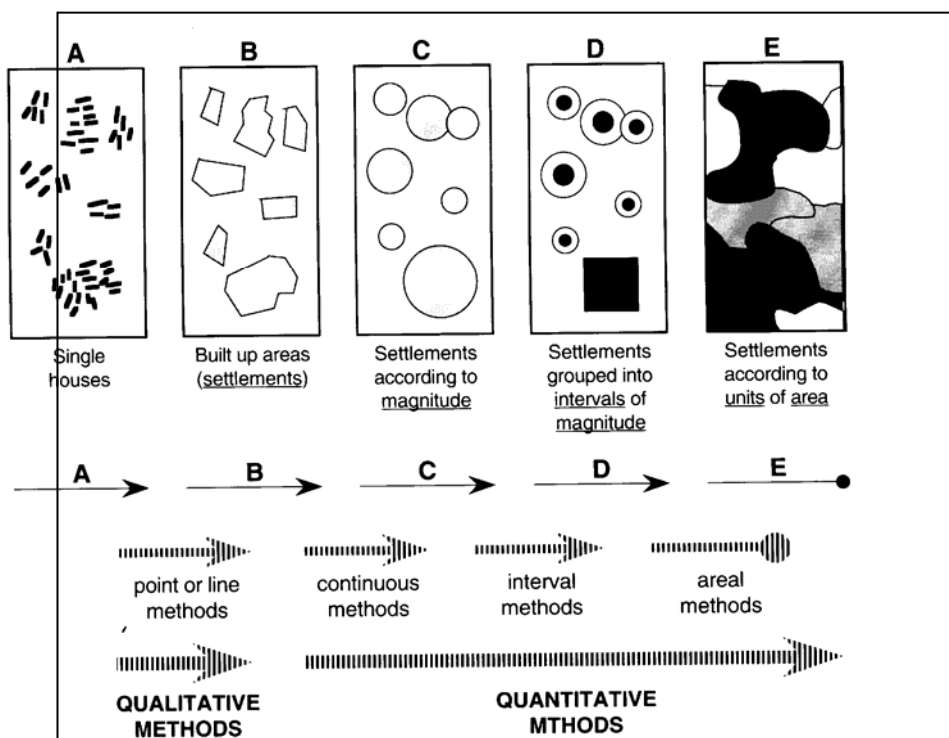
3.3. La integració de les dades en el preprocessament: l'escala d'anàlisi

La generalització possibilita tant el procés d'anàlisi com el de la representació cartogràfica. El seguiment que es fa del procés de generalització en l'entorn manual i en el digital és diferent, i la generalització en aquest segon entorn no ha pogut, de moment, emular la manera holística de controlar el procés del cartògraf experimentat. El procés de generalització que correspon a la compilació de les dades originals, amb vista a la integració de les dades difícilment aporta les claus per decidir sobre ha de ser l'escala d'anàlisi.

Una hipòtesi de treball consisteix a considerar que el procés de compilació de la cartografia manual és equivalent al del preprocessament en el model de les dades, on s'ajusta i es tria una vista del model de les dades, en funció d'una escala d'homogeneïtzació, per integrar-lo o federar-lo amb vista a l'anàlisi.

Es pot considerar que l'escala, com a escala mètrica, no té sentit en el nou entorn, però sí que en té l'escala que utilitzem al nivell més alt, al nivell conceptual, en la preparació de la informació durant el preprocessament. Una crítica que s'ha fet a la cartografia tradicional és la manca d'objectivitat durant la fase del procés d'elaboració de la informació, especialment en la simbolització de les dades, en què el cartògraf ha tendit, a través de la manipulació de unes formes, a deformar a la seva mida la realitat, i ajudat per les possibilitats de la codificació, fer "quadrar" aquest procés. La simbolització és clau en el procés cartogràfic i ha permès "embastar" diferents nivells d'anàlisi, corresponents a escales conceptuals diferents, sota el paraigües de l'objectivitat que dona el concepte d'escala mètrica i actualment, de la resolució, d'objectes digitals cartogràfics diferents. Aquesta possibilitat de discriminació i exageració que té la

simbolització, es exposat i resolt per Rastajski (McMaster, Shea 1992:) amb el model que anomena “catastrophical approach”.



(Font: McMaster, Shea 1992)

De les principals necessitats per part dels SIG durant el procés de generalització (Lagrange, Ruas 1994:1102) en la integració de dades de la visualització de dades, anàlisi de dades i derivació de base de dades, ens centrarem en la integració de dades, ja que és la que porta a plantejar el concepte d'escala. Això és especialment cert en l'actualitat amb bases de dades molt disperses, que fan necessari que poder-les integrar i per fer-ho ens hem de referir a una escala o resolució comuna.

La diversitat de fonts de les que provenen les dades i de la varietat de la seva estructura, fa difícil la integració de dades procedents del que anomenem ciències de la terra i les procedents de les ciències socials. Podem dir que, a trets generals les dades provinents de l'entorn físic són les que provenen de sensors remots i de mapes temàtics, mentre que les dades provinents de l'entorn de les persones correspon a una informació que generalment prové de cartografia ja existent, com són els límits administratius o de dades que estan associades a dades socio-econòmiques, que són dades espacials de

diferent tractament per la seva mateixa naturalesa, com pot ser tot el tipus d'informació associada a un zip-code o codi postal.

La recollida i la representació no es innòcua en el procés de comprensió final, els Sistemes d'Informació Geogràfica es nodreixen tant de dades geogràfiques espacials com de dades no espacials. Cal replantejar-se quina és la naturalesa de les dades Martin (Martin 1991:45) per poder saber quin ha de ser el tractament, per aconseguir un objectiu determinat.

De cara a la necessària integració de dades cal tenir en compte que, estem tractant dades procedents de fonts diverses com pot ésser la percepció remota o els mapes, per poder-les integrar en l'anàlisi geogràfica. Això fa que haguem de tenir en compte, que les dades satèl·lit difereixen en alguns aspectes de la resta: tenen una alta correcció posicional, una resolució espacial i temporal elevada, però en canvi un baix nivell d'abstracció humana o d'interpretació, podríem afegir que estan mancades d'escala. En canvi, els mapes representen fenòmens seleccionats de manera molt abstracta i generalitzada.

I es pot considerar que per la seva comprensió és una bona guia per l'anàlisi quan s'integren dades remotes tenir en compte el context espacial. De cara a l'anàlisi aquests dos tipus de dades difereixen a molts nivells, i necessitem de estructures múltiples de les dades (multiple data structures) que permetin la seva utilització conjunta i per el moment cap SIG ho permet ([Davis F. W., 1991 #498] :197).

La integració de les dades és el procés de fer compatibles diferents conjunt de dades; i que requereix que ens plantegem al fer-ho un seguit de qüestions que Flowerdew resumeix en les següents preguntes: ¿Quin (what) tipus de dades?, ¿A quin lloc (where) es refereixen les dades? ¿a quin moment (when) es refereixen les dades? ¿Com són d'acurades les dades? (Flowerdew 1991:375).

Dels principals aspectes que implica la integració, una primera qüestió fa referència al tipus de mesura o escalat de les dades hem utilitzat; la presència o absència d'un element; i a la categoria assignada. En segon lloc, es important de cara a la integració, el fet de que la observació es refereixi a un punt, una línia o una àrea, i quin és el sistema de referència que s'ha utilitzat. En tercer lloc hem de tenir en compte si les dades que hem d'integrar pertanyen a un moment concret o a un espai de temps i en

quart lloc, hem de conèixer tots els possibles errors de mesura, posició, imprecisió i en l'estimació.

3.3.1. La manipulació de les dades en el DLM: DLM-DLM/DKM

Una base de dades espacial seguint l'escola de Hannover, es una part de la realitat, i es caracteritza per el contingut temàtic específic i el paràmetre associat de grau de correcció que posseeix. El paràmetre de grau de correcció ens indica el nivell de detall o la mesura del màxim error de la base de dades i és desitjable que sigui igual pels diferents temes d'una base de dades. Aquest paràmetre es important en la reducció del contingut d'un model de dades en un altre de menor nivell de detall.

Com a usuaris d'un SIG seria útil disposar d'una sèrie d'eines apropiades pel tractament analític dels diferents tipus de dades geogràfiques. Per exemple, el element gràfic corresponent a punt es pot agregar com a àrea, i les àrees venir representades per un punt de referència, com es en l'exemple de gran part de les base de dades socioeconòmiques; les línies es poden agregar com àrees, com per exemple en el tractament de diferents trams de riu; i les dades d'àrees es poden convertir en superfície i els valors de superfícies estimar-se com punts i àrees. Gatrell (Gatrell 1991:). La utilització de superfície, es especialment interessant degut a la possibilitat d'integració de dades incompatibles que ofereix; cada un dels valors de cada unitat d'àrea es pot considerar com un valor de la nova superfície. Aquesta idea va ésser suggerida per Tobler (1974) i Martin la recull (1989) (cit a Flowerdew 1991:384), ja que es molt útil per poder ajusta una superfície al valor del centroid de dels polígons, i és especialment útil en les aplicacions en que es vol preparar les dades a través de una codificació que permeti la integració de dades de diferent tipus, especialment en aplicacions de tipus socioeconòmic.

En la integració de dades que provenen de diferents fonts, hem de tenir en compte: que els diferents formats de les dades publicades per diferents organismes tenen estàndards moltes vegades diferents; que poden tenir diferents projeccions, i en tercer lloc ens trobem amb el problema de no conèixer l'escala.



Table 25.1 A simple typology of some spatial analysis methods

Type of geographical data	Methods of analysis
Point	Nearest neighbour Quadrat methods
Line	Network analysis and graph theoretic methods Fractal dimension Edge detection
Area	Shape measures Spatial autocorrelation Spatial regression Regionalization Spatial interaction Location-allocation modelling
Surface	Image processing Bayesian mapping

(Font: Openshaw, 1991 #337]:390)

3.3.1.1. La manipulació de les dades en el model cartogràfic : la simbolització

En la cartografia tradicional, el cartògraf manipula la realitat a través de les representacions de les dades de manera simbòlica; això ho fa tenint en compte els objectius del mapa. Aquests objectius son en el cas de la cartografia tradicional la representació de la informació que ha de ser analitzada visualment, per el que rep les dades, o sigui el lector del mapa; es desitja doncs una eficàcia en la comunicació per el posterior anàlisi.

La generalització en la cartografia tradicional, té com a objectiu una reducció de la informació; però degut a la dificultat en la basada en les representacions cartogràfiques, i es a partir d'un canvi en l'escala de mesura i per tant en el tipus de dades i una vegada determinats els aspectes específics de la sèrie de dades a representar, el cartògraf selecciona entre els elements gràfics el més apropiat apropiats per la construcció del símbols necessaris per codificar la sèrie de dades. Hi ha combinacions però que no són factibles, i d'entre las possibles, unes son son preferibles sobre les altres.



TABLA 12.1 La problemática de la simbolización

Tipo de dato	Escala de medida			
	Nominal	Ordinal	Intervalo	De índice
Puntual	<i>tono</i> <i>forma</i> <i>orientación</i> <i>espaciado</i>	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>
Lineal	<i>tono</i> <i>forma</i> <i>espaciado</i> <i>orientación</i>	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>
Zonal	<i>tono</i> <i>espaciado</i> <i>forma</i> <i>orientación</i>	X	X	X
Volumétrico	X	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>	<i>valor</i> <i>tamaño</i>

Los elementos gráficos en cursiva son de utilidad primordial, X indica que no existe elemento gráfico apropiado.

(Font: Robinson et al. 1987:281)

De les variables diferencials dels elements gràfics espacials, el tò, l'orientació i la forma són útils per representar l'escala nominal de mesura amb el tipus de dades gràfiques corresponents el punt, la línia i l'àrea. En la cartografia tradicional el tipus de correspondència existent entre les dades puntuals, lineals o zonals se'ls fa correspondre, en la representació cartogràfica els símbols de punt, línia i àrea. I aquestes a la vegada corresponen a les dades de les variables geogràfiques recollides en base a les variacions dels fenòmens geogràfics.

El conjunt de les dades geogràfiques es poden simbolitzar basant-nos en dos aspectes: l'escala de medició i el tipus de dades. La generalització cartogràfica al simbolitzar ho fa a dos nivells:

- 1) fent un canvi d'escala de mesura; a partir del conjunt de dades originals
- 2) fent un canvi en el tipus de dades; això suposa mantenir el sentit en l'escalat de les dades de nominal a ordinal, i d'interval a índex

Quan volem integrar dades en el model cartogràfic, hem de seguir les regles desenvolupades en la cartografia que permeten la creació de nous objectes a partir de la combinació de els elements gràfics corresponents a la seva representació cartogràfica.

	Strukturtypus	punkt-, linien- oder flächenbezogen	Niveau der Information
1	Kontinua	flächenbezogen	quantitative Werte
2	Flächenstufen	flächenbezogen	geordnete Stufen
3	Flächenmosaike	flächenbezogen	qualitative Unterscheidungen
4	Flächendiagramme	flächenbezogen	je quantitativ, geordnet oder qualitativ
5	Standorte	punktbezogen	qualitative Unterscheidungen, ev. auch geordnete Werte
6	Streuungen	punktbezogen	Einheits-Mengenwerte
7	Punktogramme	punktbezogen	je quantitativ, geordnet oder qualitativ
8	Liniennetze	linienbezogen	qualitative Unterscheidungen, gestufte Werte
9	Liniendiagramme	linienbezogen	je quantitativ, geordnet oder qualitativ
1 _t bis 9 _t	dynamische Vorgänge, zusätzlich bei allen obigen Typen möglich	punkt-, linien- oder flächenbezogen	alle drei Informationsniveaus

(Font: Spiess 1990:64)

Com molt be assenyala Robinsón, els elements gràfics connoten certes escales de mesura i varien en quant a la seva utilitat de representació dels diferents tipus de dades; també assenyala que, la relació entre els aspectes seleccionats de la realitat, i els elements gràfics primaris, i és decisió del cartògraf: decidir quins aspectes seleccionar, quins elements gràfics primaris se'ls ha de fer correspondre, ja que la comprensió del mapa rau precisament en això.

"Si la simbolización se utiliza de manera correcta, el diseño bueno o malo no alterará gravemente la utilidad del mapa"([Robinsón, 1984 #476]:279)

Una altre possibilitat, menys utilitzada, es el canvi de la dimensionalitat dels objectes cartogràfics, el mateix Robinsón, cartògraf experimentat, assenyala que les abstraccions derivades d'un canvi de dimensions, encara que poden desconcertar a vegades, son de gran utilitat en la comprensió de les característiques de la distribució

espacial o en la comparació d'aquestes distribucions. Tal com fa Muller, també aconsella que el cartògraf en el futur explori aquesta possibilitat; ja que l'ajut de l'ordinador faciliten la manipulació prèvia en el model de les dades, difícil de realitzar anteriorment (Robinson et al. 1984:279).

3.3.2 Escala conceptual d'homogenització per l'anàlisi

Un dels conflictes existents en el mapa analògic tradicional, era la relació entre el grau de correcció mètrica i de representació visual, això tendeix a desaparèixer amb la tecnologia electrònica, o al menys a perdre importància.

"The history of terrain representation over the past three centuries is one of the competition between the display of accurate data about elevations, slopes and aspects of the ground surface of the ground, in a given area of interest."

L'efecte del canvi tecnològic permet precisament que el grau de correcció ja no sigui la precisió mètrica, sinó que permet altres tipus de precisió com pot ser el de la precisió semàntica, dels fenòmens que s'han de representar.

"...cartographers can concentrate on the aproprament al valor real and precision of the atribution and also on the realistic nature of the visual display..."

En aquest context la importància de l'escala mètrica en el mapa tradicional, i el control de l'escala i la seva concreció és de cabdal importància pel disseny del mapa. Morrison tot i haver-se format en la cartografia tradicional, veu la necessitat en haver de canviar el concepte que guia el disseny del mapa de l'escala mètrica, i ho fa donant aquesta importància a la "resolució". Aquest element tindria l'avantatge de formar part de els elements de teledetecció i per tant que representen un element comú en les necessitats d'integració de les diferents dades.

"Using electronic technology scale of the visualization does not have to be static [...] generalization becomes situation or query specific. The controlling factor is the resolution of the digitized feature and its attributes contained in the database. Thus, resolution of data becomes of major concern to the producer/cartographer, and information about the resolution of the data must be contained in metadata accompanying each feature so that user/cartographer can determine appropriate uses. Scale of use becomes a variable that the user/cartographer assumes control over, and can vary at will. " ([Morrison, 1994 #328]:7)

3.3.2.1 Les noves possibilitats de l'anàlisi espacial

La necessitat de conversió de les dades des de les representacions cartogràfiques a fer-ho partint de les dades originals per el seu anàlisi, es deu en gran part a la necessitat d'integrar dades procedents de l'entorn físic i de l'entorn socioeconòmic, juntament amb la possibilitat de disposar d'una tecnologia que permet manipular grans quantitats de dades.

Els SIG tenen en la cartografia però també en la geografia quantitativa la seva fonamentació, en el moment actual. Així ho esposa Openshaw cita a Hagerstrand cercant una definició adequada d'anàlisi espacial

"... to a small degree the recent quantitative analysis in geography represents a study in depth of the patterns of points, lines, areas and surfaces depicted on maps of some sort of defines by coordinates in two-or-three dimensional space..."

I a continuació evidencia com:

"...it is important of increasing the functionality of gis by providing a link between the essentially cartographic domain in which the origins of GIS lie and key areas of applied quantitative, statistical and mathematical analysis, and modeling, of interest to many users of GIS. However in seeking to meet these objectives the recommended technology has to be capable of coping with the peculiarly complex nature of the spatial data..." (Openshaw 1991:390).

En aquesta línia es trobaria també Bailey (1994:14) que diferencia entre el que es considera l'anàlisi espacial i el més específic de l'estadística espacial. Degut al fet de desenvolupar-se l'anàlisi espacial en la geografia quantitativa i estadística dels '50, l'anàlisi espacial s'ha basat en la aplicació de mètodes estadístics a les dades espacials (Berry i Marble,1968) i models matemàtics (Taylor, 1977). Però encara ara la majoria de SIG han integrat recentment les possibilitats d'anàlisi espacial, Openshaw (Openshaw 1991:390) posa en evidència com dels 1000 comandaments, pocs són de anàlisi espacial i son encara menys si eliminem els que pertanyen a la manipulació de dades: que no són realment d'anàlisi sinó que són descriptius o cartogràfics com son els "bufferings" superposició de capes o de consulta.

Per tal d'integrar les dades per poder fer anàlisi espacial, hem de en compte que: la possibilitat de millorar els mètodes tradicionals de visualització, d'exploració i de representació de les dades espacials; així com la necessitat d'ajudar als usuaris dels SIG

a entendre i millorar els mètodes d'anàlisi espacial que estan utilitzant. Els sistemes d'informació geogràfica poden introduir funcions d'anàlisi espacial, i a la vegada millorar aquestes, a través de per exemple el reconeixement de patrons espacials, com es fa actualment en teledetecció.

“that geographers need to play more of a role in ensuring “the timely and accurate usage of sophisticated spatial analysis tools” in a gis environment”. (Kelly 1994:65

3.3.2.2. L'escala

Si utilitzem el terme escala tal com s'entén en la cartografia tradicional, segurament pensarem en l'escala mètrica, ja que la definició d'escala en la cartografia tradicional fa referència a la relació existent entre la distància en la realitat i en la seva corresponent distància en el mapa i dona per descomptat de que el fenomen es evident en el món real. Així doncs, l'equivalent del concepte d'escala per gran part dels geògrafs seria el que ve donat per la resolució, normalment referida en el seu vessant espacial, no espectral ni temporal. Hem de tenir en compte que parlem de la escala mètrica sense entrar en el problema que comporten les projeccions i que fan que sigui necessari a nivell mètric parlar del factor d'escala.

Però això no és així, ja que en funció del nostre interès captem la realitat de manera diversa agregant o desagregant elements de la realitat. La fase de selecció de la realitat la fem en base a les relacions de certs fenòmens entre ells, i és el que integrem en un model de dades. El pas d'un fenomen del món real o “entitat”, a la seva conversió en un “objecte digital” no és doncs factible si ens guiem només per la seva resolució mètrica i no ho fem a un nivell més alt d'abstracció.

Existeix una acomodació a plantejar certs fenòmens a determinades escales conceptuals i estem acostumats a detectar un tipus de fenòmens a certes escales, que ens permeten acollir-nos a un marc d'anàlisi conegut. O sigui, representem diferents fets a escales determinades i no a d'altres. De la mateixa manera que quan llegim el diari el nostre llindar de susceptibilitat i comprensió davant del problema que s'exposa es diferent en les pàgines internacionals que en les locals.

Langran Langran 1992 suggereix un exercici de comprovació, consistent en suprimir l'element objecte d'estudi: per sostreure'ns del pes de la inèrcia de la tradició i provar a canviar una variable en el procés per veure com repercuteix sobre la resta de variables.

És evident que l'escala mètrica per si sola, no permet assolir la gran quantitat de

tasques que tradicionalment se l'hi adjudiquen, en el desenvolupament en el desenvolupament del procés cartogràfic.

La integració la fem a partir de l'escala conceptual, i no de l'escala mètrica; tot i que les entitats espacials ja comporten el seu propi llinyar de resolució, no és cert que a qualsevol escala sigui extrapolable la resolució. En altres paraules, encara que puguem convertir un fenomen a través de la conversió de la seva resolució a una altra resolució és molt possible que aquest fenomen no tingui "sentit" fora del context que es redueix al que Muller exemplifica com el rang d'escala "permeses" per un determinat fenomen. Quan compilem les dades, ja fem una tria dels fenòmens del món real que tenen "sentit" a una escala conceptual determinada, i l'escala mètrica és només una característica, com ho és la resolució espacial d'un fenomen, que integrem a la base de dades, i que es relaciona amb certs fenòmens dins d'una mateixa escala.

Escala i resolució podrien considerar-se a cada extrem de les maneres de la conceptualització de l'espai, i partint d'una escala i un nivell de detall del fenomen que volem analitzar podem arribar a concretar la resolució espacial a la que podem representar aquest fenomen. No a l'inrevés; ja que a partir de l'escala ens fem idea de a quin nivell conceptual ens movem i quin tipus d'anàlisi podem fer sense saber si la resolució serà l'adequada. La resolució ens permet fer el contrari, analitzar un fenomen sense saber si el resultat serà significatiu al nivell conceptual volgut.

En el procés de selecció, simplificació tant geomètrica com conceptual que comporta una reducció de la informació el que Günreich anomena generalització de l'objecte (object generalization) i que constitueix el model primari o mapa base topogràfic (DLM). A partir d'aquest model hem de poder derivar altres models, bé de dades o bé cartogràfic, i la informació continguda en la base de dades les dona la resolució espacial o mètrica; però les series cartogràfiques son "estanques", en la transició dels objectes i fenòmens corresponents entre elles; i aquesta correspondència entre la representació d'un fenomen i el rang d'escala en que es pot moure la seva representació, no és deduïble de la informació en les bases de dades, actualment.

Ja s'ha vist en el capítol 2 que en l'entorn digital s'amplien les necessitats de la comunicació cartogràfica a la manipulació de la base de dades. De les necessitats existents en el procés tradicional, el canvi d'escala és un concepte clau en la conducció

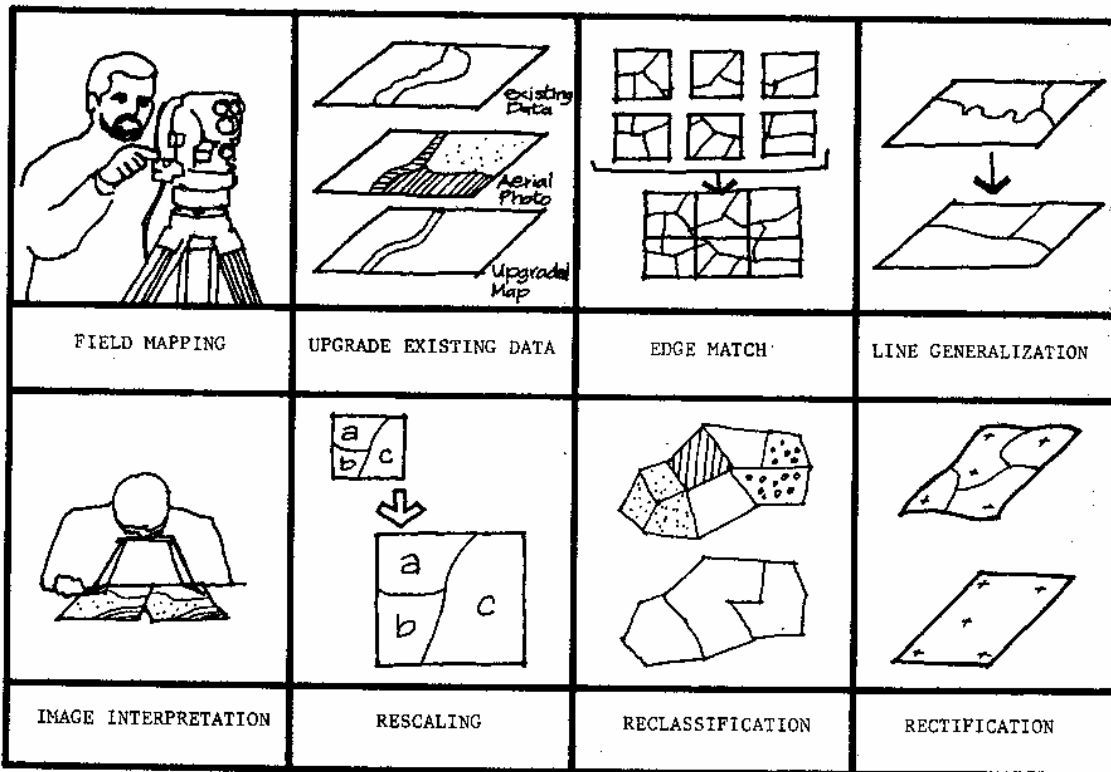
del procés. En el nou entorn, però, el concepte d'escala tradicional desapareix i segons molts autors ens hem de referir a resolució. Aquesta possibilitat té sobretot sentit en el camp de la percepció remota, però es evident que escala i resolució estan en els extrems d'un eix que permet relacionar les dimensions de les entitats i el sentit de representació d'aquestes a unes resolucions determinades.

Sembla possible afirmar que l'escala ens porta a unes resolucions de representació però, a l'inrevés, això no es tant clar. Per exemple, quan nosaltres volem representar un fenomen, triem un sensor remot d'una resolució adequada, per l'anàlisi d'un determinat fenomen geogràfic. El que no és tant automàtic, però, és saber, a partir d'una resolució determinada, quins fenòmens són aptes per poder ser representats i a quina escala ho haurem de fer.

Un problema que s'ens presenta en l'automatització de la generalització es la de poder determinar en quin moment assolim l'objectiu desitjat; Tobler suggereix que un fenomen s'entén a l'escala a la que es maximitza la seva variança (Tobler 1988:).

Tots aquests problemes estan lligats a la manca, ja no d'una escala mètrica de representació cartogràfica, sinó d'una escala conceptual explícita en el procés de compilació de les dades.

Tant en el procés de compilació manual com en el digital, un primer pas consisteix en avaluar els materials procedents de les diverses fonts, per establir quin tipus d'informació resultant podem generar. Aquest procés ens han de permetre poder decidir, en funció dels requisits del mapa final, quines transformacions a les dades originals s'hem de realitzar. Una norma del procés de compilació de la cartografia manual es la de derivar dades de menor resolució d'un model de major resolució, amb una cartografia original que s'aconsella tingui una relació de 150% a 200% d'ampliació. En les dades digitals és aconsellable poder utilitzar les dades provinents del model primari, o sigui que tendim a utilitzar dades originals, com fa la teledetecció. Aquesta utilitza tècniques de visualització de les dades i de tractament de la imatge, però l'anàlisi s'efectua sempre no sobre les dades que estem visualitzant sino sobre les dades originals.



THEMATIC MAPPING TECHNIQUES

(Font: Dangermond 1990)

Ara be, cada vegada més sovint s'utilitzen dades originals pel processament, i el principal problema que sorgeix en la utilització de les base de dades originals, és el fet de no tenir escales associades a les entitats enregistrades. Sí que tenim una idea procedent del nostre coneixement, dins de quin rang d'escales un fenomen es troba, ja que ens basem en una tradició de representació i anàlisi previ representat en mapes temàtics.

En les metadades associades a una base de dades disposem de la resolució d'aquesta, però no de la escala conceptual o de percepció. De manera ideal Muller assenyala que hauríem de poder variar la escala en funció del nivell de precisió desitjat.

"One might ignore the problem of representation and vary the level of precision (which amounts to a scale variation) at which spatial objects,..., are encoded in the database"

I a continuació fa evident que el poder comptar amb el llinatge o tracprocés de tractament que han sofert les dades ens indica el grau de correcció i el nivell de

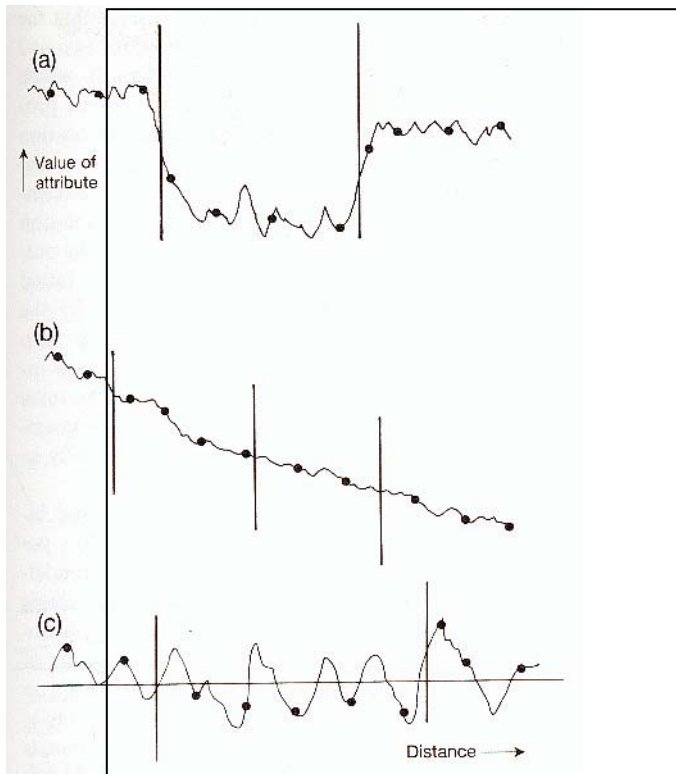
resolució i per tant d'agregació d'un fenomen espacial.

Asking the scale of a database is in fact making a query about the lineage of the database" (NCGIA). This lineage is not the same for all objects, some being better documented than others". (Muller 1991:459)

El sistema d'estructura de les dades que utilitzem es important ja que ens fa entrar en dos maneres conceptualment molt diferents de considerar els fets geogràfics. De les dos grans estructures de dades, vector i raster, la primera estaria associada als fenòmens o entitats que representem amb línies o fronteres, mentre que el segon, ens permet representar els fenòmens com a objectes considerats de manera més probabilística. Identificar un element pel seu contingut, o sigui en base a la seva variança, enlloc de fer-ho pel seu continent o fronteres de l'element, seria equiparable a les veure les diferències existents entre el model de dades de tipus raster i el de tipus vectorial. Aquesta distinció es important a la hora de aplicar els diferents operadors com molt be sintetitza McMaster (McMaster, Shea 1992).

Intentar trobar regions homogènies i patrons espacials ha estat sempre fonamental en la investigació geogràfica i estadística. El preprocessament de les dades permet precisament contribuir al seu coneixement a "veure el que no es veu" no és evident ([Buttenfield B P, 1991 #478] :427). Ja be sigui amb l'ajut de la tecnologia actual, que permet extraure informació visual espacial, o bé amb el tractament estadístic o numèric de les dades originals.

Trobar regions homogènies depén de la consideració de la naturalesa dels seus límits ja que considerem que les entitats que observem al mon real són representables amb una línia clara sobre el mapa. Això es així per les delimitacions administratives però no pels fenòmens "naturals", que els podríem considerar "fuzzy" o poc definits. Ens pot ajudar en aquesta discussió la contribució de Burrough de que les variacions dels fenòmens naturals representats per fronteres són difícilment definibles com abruptes o suaus (Burrough 1986:120). Però aquesta característica depen de l'escala, la caracterització d'un fenomen a dos escales diferents, és diferent. El mateix passa amb altres característiques dels objectes cartogràfics i en les fotos a on el concepte de textura varia en funció de l'escala.



(Font: Burrough 1986:121)

Podríem, doncs, considerar la possibilitat de que definir límits d'arees homogènies a partir d'una consideració estadística, traçant les fronteres allà on el valor de les propietats canvien amb la distància. Per tant, el concepte de frontera estaria lligat al lloc on la taxa mitjana de canvi del valor d'una propietat, és màxim. De la fig anterior podriem doncs distingir fronteres abruptes resultants de canvis abruptes i “large changes in the value of critical properties over short distances” canvis grans en els valors corresponents de propietats importants en arees petites. Unes altres fronteres que es podrien considerar zones contínues i que són més difícils de poder dividir poden ser les que corresponen, per exemple, a gradients climàtics. L'últim tipus correspon quan dues observacions son classificades diferentment, **perque el valor del criteri de discriminació, consisteix en pols oposats d'una classe** “occurs solely as a result of two observations being classified differently, simply because the value of a discriminating criterion lies on the opposite sides of a class, as distinct from a landscape barrier”. Es un tipus de frontera que no té cap mena de sentit en el context de recerca de el tipus de

trobar quina és la petita variació dominant.

Un problema que sorgeix es que el grau d'abrupta d'una frontera llandant entre dos objectes espacials, varia en la seva llargada i l'altre és que la localització i llargada d'una frontera és funció de la seva escala. Això, demostrat per Richardsón (cit. Burrough 1986:122), és degut a que a mida que l'escala del mapa s'incrementa, es poden delinear més elements, que no es podien abans, i Richardson utilitza una línia de costa per demostrar com creix de manera logarítmica amb l'increment de l'escala del mapa. Això s'ha vingut desenvolupant en la teoria dels fractals teoria exposada per Mandelbrot (1962) que reconeix que en la major part de fenòmens naturals, la quantitat de detall (resolvable) és una funció de l'escala, i incrementar l'escala del mapa no implica un increment absolut de la precisió sinó que revela variacions altrament imperceptibles. La teoria dels fractals ha aportat dos idees que han tingut els seus seguidors en el camp de la cartografia basant-se en: la idea de (self-similarity), o sigui que la manera de variar un objecte espacial entre diferents escales té un comportament similar i que ve donat en la seva propia morfologia; i en segon lloc la característica de tenir una dimensió fractal. Zonneveld (1973) rebut la idea de que les variacions de les formes del paisatge al llarg d'uns quants metres siguin equiparables o similars estadísticament a d'altres al llarg de centenars de milers de metres quan les formes son transformades per un paràmatre d'escalat simple.

En els estudis de fractals no podem mirar les fronteres com a absolutes ja que són compromisos i la seva llargada es una complexa funció de la variació de la propietat que es cartografia, de la escala i del mètode de cartografiar. Han permès pensar en noves maneres de representar les variacions dels fenòmens naturals.

La escala i les "escales incompatibles"

Podríem considerar que les entitats estan relacionades a l'escala conceptual i els objectes espacials a la resolució, i hi ha un moment en que el canvi de resolució ens mostra fenòmens diferents. (Es meu)

es una idea de la nuria i crec que es potser millor si ho definim com entitats que els definim en espais incompatibles

En el quadre de Uninayar [Uninayar, 1988 #512]:

Representem un fenòmen del mon real a través d'un objecte espacial i podem fer-l'hi

correspondre a diferents escales diferents representacions cartogràfiques, i correspon al que s'anomenen "representacions múltiples". Degut a la falta de continuïtat en la reconversió de la representació gràfica de les representacions múltiples i la impossibilitat de manteniment del "sentit" conceptual, en aquestes és un dels problemes que centren l'interès en la generalització cartogràfica actualment. Això té com a conseqüència, que un objecte no el podem analitzar a totes les escales mètriques ja que probablement perdi "sentit" conceptual.

Com molt bé insinuen Ruas i Lagrange ([Ruas A., 1995 #488]:73) en "generalization: precision variation or change of perception level" un mapa correspon a una visió (view) de l'espai geogràfic i la percepció d'aquest espai canvia amb l'escala. Quan l'escala varia, l'espai no es pot (depict) de la mateixa manera, utilitzant els mateixos objectes. I a la inversa, un canvi en la percepció indueix a un canvi en el nivell d'escala. La generalització s'hauria d'entendre com un procés que realitza o fa (perform) un canvi en el nivell de percepció de les dades geogràfiques, i a on la precisió i els canvis en la geometria no són més que conseqüències d'aquest procés.

Ruas i Lagrange plantegen el que anomenen l'espai fix i el variable que en el cas del mapa topogràfic correspon a l'espai mètric i al temàtic i es innegable la eficiència que aquest representa, si tenim en comte la quantitat d'informació que contenen el mapa topogràfic, una bona entente entre els objectes representats i les seves relacions espacials.

L'espai de conceptualització i l'escala

No es fins recentment que l'estudi de l'escala associada al procés cognitiu és prioritari dins la recerca dels SIG. La NCGIA (National Center for Geographic Information and Analysis) que identificà com a tasca prioritària, durant els anys 80'centrar-se en les estratègies de manipulació de la informació geogràfica i la seva anàlisi; en el programa de recerca pel segle XXI, en l'anomenat projecte Varenius dona prioritat a tres àrees que es centren principalment en el desenvolupament de processos cognitius. En primer lloc, pretén aprofundir en el coneixement de les àrees del saber a on es fa la formalització dels conceptes geogràfics, que permeti avançar en el desenvolupament de eines i mètodes dels fenòmens geogràfics, i entendre com es distribueixen espacialment aquests. En segon lloc, la recerca ha de permetre entendre conceptes bàsics del

coneixement geogràfic que es necessiten per produir noves tecnologies. En tercer lloc, examinar els impactes que aquestes tecnologies tenen sobre els individus, organitzacions i la societat en general.

De les tres àrees estratègiques de recerca en la Informació Geogràfica la primera d'elles són els models cognitius de l'espai geogràfic⁶. En la primera de les tres àrees estratègiques de recerca, la que correspon als models cognitius de l'espai, i en especial a l'escala, té alta prioritats des del punt de vista conceptual ja que és un concepte fonamental geogràfic i que presenta molts problemes en els sistemes digitals geogràfics. El projecte Varenius presenta l'escala i el detall en el procés de cognició de la Informació Geogràfica, i assenyala que el terme escala té múltiples referents, dimensió absoluta, dimensió relativa, resolució, granularia i detall.

Es fa palesa la necessitat de referir-se als múltiples conceptes d'escala, amb especial atenció als aspectes cognitius de l'escala per sobre de la donada a l'escala a les representacions externes. En aquesta sentit es planteja la pregunta:

"Can we identify the fundamental, invariant aspects of the concept of scale that survive the transition to the digital world?"
NCGIA et al. 1997

La resposta de Ruas i Lagrange és que la generalització s'ha de veure com un procés que ha de permetre un canvi en el nivell de percepció de les dades geogràfiques, on el canvi en la precisió i la geometria no són més que conseqüències d'aquests canvis (Ruas, Lagrange 1995:73).

La integració de dades d'estructura diferent

La integració de diferents estructures de dades principalment corresponent a vector i cel·la es necessari i la majoria de programes actualment són més flexibles que no ho han estat, són els IGIS Integrated Geographic Information Systems, sistemes capaços de processar dades vector i raster. De totes maneres un problema que es planteja en el moment d'integrar, o fer compatibles dades de diferent tipus, en bases de dades que es puguin federar en determinats moments per poder ser utilitzades, és l'escala que donem

- _____

⁶ L'interès en aquesta àrea es degut al fet de que els sistemes d'informació venen "on line" es necessari conèixer més sobre el espai cognitiu ja que de moment els usuaris han de ésser especialistes espacials.

a els nous objectes.

El sistema més senzill seria el corresponent als sistemes que tenen eines que permeten la conversió i la visualització de les dades dels dos formats. I el desitjable seria aquell sistema que mantingués indistintament representacions objecte-orientades i camp-orientades de les dades geogràfiques. En aquesta direcció es troba el grup de recerca del LUCC.

Una possibilitat seria utilitzar les tècniques d'estadística espacial en base a la variació interna dels objectes espacials representats en la base de dades. Una altre seria fer-ho en base a la seva integració a un nivell més alta d'abstracció, a nivell semàntic. Al ser els processos geogràfics són escala dependents, seria ideal poder modelitzar la informació geogràfica, com assenyala Tobler: al nivell a on la variança geogràfica es maximitza o al nivell en que els processos geogràfics s'entenen millor (cit per Müller 1991, Tobler and Mollering 1972).

Si distingim una integració de dades procedents de diferents fonts, hem de tenir en compte que a les actuals tecnologies s'han aplicat sobretot al medi físic, i no a les dades socioeconòmiques i que el tractament hauria de ser diferent encara que corresponguin a objectes digitals iguals.

“...the requirements of GIS for handling socioeconomic information are in many cases different from those concerned with phenomena in the physical environment...” (Martin 1991:28)

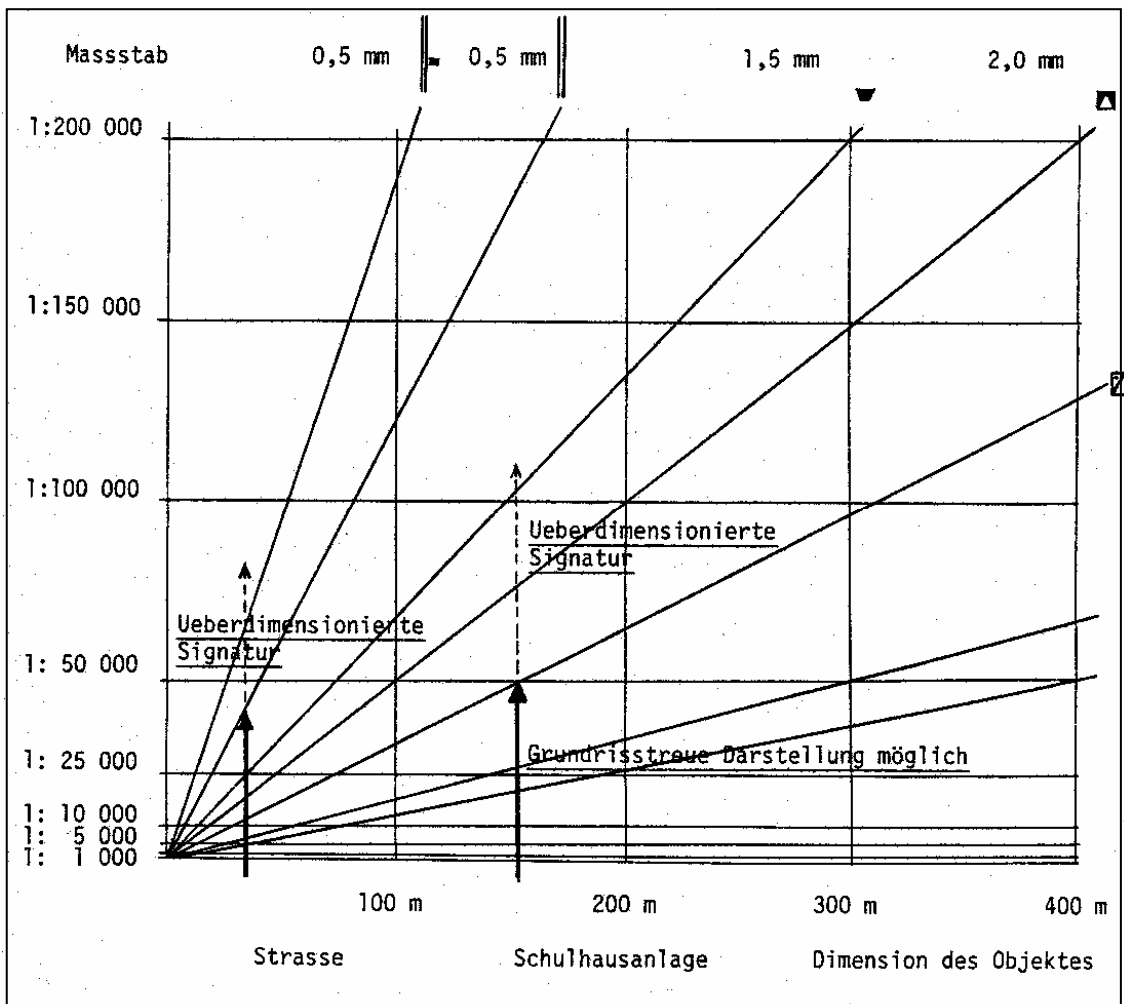
Les dades socioeconòmiques tenen uns problemes diferents com provenen de la diversitat d'organismes i de finalitats encara no totalment establertes, el que es veu en la discussió sobre la dimensió de l'àrea mínima a la que s'han de recollir les dades relacionades amb les variables socioeconòmiques. La àrea de recerca petites o “small areas”; requereixen d'un tractament estadístic espacial que tingui en compte el problema que suposa la intervenció de l'espai en un grau alt de autocorrelació en àrees properes enentre si. La determinació del tamany d'àrea mínima és l'equivalent de la resolució espacial utilitzada en el camp de les ciències de la terra.

“the variety of basic units has presented a major obstacle to the analysis of socioeconomic data, as the patterns apparent in the data may be as much due to the nature of the collection units as to the underlying phenomén, and there is no direct way of comparing data collected for differing sets of areal units.” (cita a Flowerdew and Openshaw 1987 Martin 1991:)

Una font de treball és la recerca d' una nova manera de representar els fenòmens, que permeti discretitzar els objectes des de l' alternativa que ve representada per considerar un espai de variació continu (NCGIA et al. 1997)

3.3.2.3 La resolució

La resolució espacial és un concepte que, en la cartografia tradicional no té el mateix significat que té en el moment actual, ja que està més lligat a les dades originals, i els mapes, a diferència de les fotos aèries i de les imatges satèl·lit, no tenen un nivell de resolució fix; la resolució permet obtenir informació sobre un element, la seva posició i la del seu contingut. Tobler, a partir de la definició de resolució espacial com la relació entre l'escala del mapa i la resolució geogràfica, estableix una taula de correspondències (Tobler 1988:132) on es guia per la regla de dividir el denominador de l'escala del mapa per 1, 000 per obtenir el tamany detectable; i així mateix ho fa Spiess per saber el tamany mínim representable.



(Font: Spiess 1982)

En la percepció remota la resolució ha estat un concepte fonamental; es defineix en relació a la possibilitat mètrica de detecció i ve donada en unitats lineals. El gran interès de poder integrar dades procedents de percepció remota i les procedents de mapes ha fet que el terme de resolució sigui utilitzat com el concepte que permeti la integració de dades diverses. Però alhora és prou clar que el concepte de resolució no és suficient per integrar fenòmens, ja que necessitem fer-ho en base al seu contingut geogràfic.

La resolució espacial, en les bases de dades, està determinada per la qualitat del mostreig; un mostreig més fí o desagregat, dona millors resolucions que un de més groller o més agregat. La generalització, en tendir al mínim tamany detectable d'un fenomen, representa la informació a uns nivells de resolució encara més grollers o agregats.

Hem de tenir en compte que la manera com es recullen les dades no és homogènia i que pot variar de lloc en lloc. Muller, recollint la tradició de la percepció remota de tractar la resolució com patrons espacials, considera que els objectes espacials i els seus processos tenen diferent tamany i diferents longitud d'ona. I, per tant, les dades s'han de representar utilitzant diferents resolucions espacials però sense oblidar que, en generalitzar, diferent tamany que caracteritza un fenomen de les dades originals, es pot mantenir preservant la resolució de la variança de la mostra original.

La generalització el que fa es reduir la resolució d'alguns fenòmens, que visibles a grans escales desapareixen a escales mes petites, a la vegada que en aquestes altres escales es fan evidents uns altres fenòmens. Muller posa com a exemple la divisió dels comptats americans en els quals es fa palés que diferents nivells d'agregació, permeten veure apareixer i desaparèixer fenòmens diferents. La generalització permet descobrir, a través de l'agregació en base a la reducció de resolució, l'aparició de patrons a diverses resolucions. I citant a Tobler i Mollering (1972)

"...a coarser resolution or a higher level of generalization may provide more explanation over the variance of a spatial variable than finer resolution levels..." (cit. de Muller 1991:461)

Tobler (1989) també assenyala que una interpolació (conversió de les observacions fetes sobre la retícula del mapa en una retícula uniforme lat/long) és pot fer si les diferents base de dades tenen el mateix tamany; pero si volem comparar dues retícules que difereixen en aquesta magnitud, hem de tenir en compte el principi de que la resolució d'una bases de dades pot ser només la de l'element més groller. Per tant, la interpolació no pot millorar la resolució de la bases de dades i només pot fer-ho si s'introdueix informació addicional del exterior.

"The spatial resolution of a geographical map is fairly easy to asses. This is because there is a relation between map scale and geographic resolution. This is readily obtained by observing that the smallest physical mark which the cartographer can make is about one half millimeter in size. Than make a simple rule: divide the denominator of the map scale by 1,000 to get the detectable size in meters. The resolution is one half of this amount. Of course the cartographer fudges. He makes things which are too small to detect much larger on the map because of their importance. But this cannot be done for everything so that most features less than resolution size get left off the map. This is why the spatial resolution is so critical...Should electronic database also be provided at comparable resolution intervals , or can procedures be developed which require that the data only be stored once, at a very fine resolution, with algorithms which provide the resolution appropriate for the problem solved?" (Tobler 1988:131)

Un concepte que desenvolupa és el del nivell de detall; la tria del nivell apropiat de resolució espacial és crítica, i també ho és l'elecció del sistema d'emmagatzematge de les dades geogràfiques en concret, posa com a exemple el fet que en les ciències de la terra la resolució espacial hagi estat associada normalment a l'interval de mesura que definim com adequat.

“From sampling theory it is known that the detection of a feature is only possible if the sampling rate is twice as fine as the size of the feature to be detected... This has an important implication for geographical data collection and storage; it implies that one must know the spatial size of the features in which one is interested before one starts to collect data. And one cannot expect a collection of geographic data to be suitable for all kinds of problems.” (Tobler 1988:134)

En la fotografia aèria és on es veu la relació entre detectabilitat i resolució; comparant una escena digital procedent d'un sensor remot a diferents nivells de resolució. Per tant Tobler ens indica una relació entre escala procedent de mapa digitalitzat o escanejat, resolució i detecció d'un fenomen.

“The definition given (here) for spatial resolution enables one to make the calculation for point, line, polygon or raster (gis) and for categorical, scalar, vector or tensor data. Comparisons between gis are most easily made when they have similar resolutions and resolutions-variances.” (Tobler 1988:135)

La resolució en la base de dades actua com a substitució de l'escala, ja que indica el nivell de detall a la que es pot analitzar un fenomen. El concepte de resolució es l'utilitzat en percepció remota i és aquesta resolució espacial a la que fa referència Langran, juntament amb el concepte de resolució temporal. En aquest cas, el concepte de resolució temporal difereix en cartografia i en teledetecció.

“...spatial resolution has been defined as the minimum difference between two independently measured or computed values that can be distinguished by the measurement or analytical method being considered or used. This definition applies well to temporal resolution...” (Langran 1992:84)

McMaster, adverteix també, de la seva banda, de la possible confusió entre escala i resolució. L'escala fa referència a l'àrea que es cobreix en l'anàlisi espacial d'un estat, una regió, o unes zones urbanes, i des de les considerades petites a les grans escales. La resolució fa referència a les unitats d'enumeració que s'utilitzen en l'anàlisi espacial.

Per tant, escala, resolució i mètodes d'anàlisi espacial estan lligats estretament, ja que un nivell groller correspondria a l'anàlisi espacial de petita escala i a l'inrevés.

McMaster, també mostra com les diferències en escales/resolucions de les àrees d'estudi i les dades geodemogràfiques, repercuteixen en les metodologies espacials i estadístiques utilitzades; com en l'exemple que proporciona, citant Cutter (1996), en un treball sobre localització de materials perillosos en el que les diferents unitats espacials d'anàlisi (escala i resolució) es produeixen a diferents nivells de correlació; en el cas d'unitats més grans d'anàlisi comportarien correlacions més fortes.

McMaster assenyala com el problema metodològic de la generalització ve limitat per les dades de què disposem i la mesura utilitzada, l'escala i/o resolució d'anàlisi i el tipus d'anàlisi estadística o espacial utilitzada. I fa una tipologia de problemes que diferencia en primer lloc en base a les dades, i la seva mesura; en segon lloc, les qüestions d'escala i resolució; i en tercer lloc els mètodes d'anàlisi i dona les possibles solucions corresponents als tres problemes ([McMaster, 1997 #238]:178). És evident que la majoria de problemes geogràfics es mouen en una franja d'escales i resolucions possibles. La selecció de la resolució s'ha de realitzar en base a l'escala d'anàlisi, així com al mètode d'anàlisi que s'utilitzarà.

"It has been well documented that the results of a given spatial analysis change with the resolution and definition of spatial units-frequently dubbed the modifiable areal unit problem- and recent gis-based assessments of environmental equity have shown that general conclusions vary dramatically as scale and resolution are altered..." (McMaster et al. 1997 :178)

Com que gran part de les dades geogràfiques han estat generades i emmagatzemades com a mapes temàtics, les dades originals no han estat el punt de partida del processament. En l'actualitat, els serveis cartogràfics ofereixen aquestes dades en format digital, i ja és indubtable la nostra responsabilitat en escollir un nivell de resolució adequat a un nivell adequat d'anàlisi o de representació.

Degut a que tant la generalització estadística com la cartogràfica afecta el grau de correcció de les bases de dades espacials, és important mantenir al llarg del procés de generalització, un control de la variabilitat espacial.

Goodchild defineix la resolució d'un producte digital, les dades digitals, en termes de la

seva resolució espacial com:

“as one half of the unit of the least significant digit”.

Seria desitjable, per tant, que la precisió coincidís amb el grau de correcció per al·ludir a la mesura aritmètica (número de decimals). La definició donada per Goodchild de precisió és la següent:

“...degree of detail in the reporting of a measurement...”[cit Muller, 1993 #515]

Defineix Goodchild [Goodchild, 1989 #296]:33, el grau de correcció a un nivell més alt o complex que correpon a una idea més ampla que el de la precisió numèrica :

“...relationship between a measurement and the reality which it purports to represent...” ([cit Muller, 1993 #515])

És a dir, la magnitud de la diferència entre el valor obtingut i el valor verdader, i per tant, el grau de correcció depèn de com definim la “veritat real”.

Les dades poden entrar-se a diferents escales, però la informació sobre l’escala original a la que s’ha entrat el document no es presava, encara que sabem que és un indicador del grau de correcció del conjunt de dades.

Tobler fa referència al fet de fer una reorganització d’una mostra si desitgem fer una conversió entre dos sistemes que no tenen el mateix ordre de magnitud, si un té una resolució de 10 metres i es troba sobre una retícula distanciada 10 metres entre els seus punts de la quadrícula ortogonal i un altre a 100 m respectivament, hauriem de tenir en compte que les resolucions seran de 10 m i 100 m i en convertir-les, hauriam de tenir en compte el principi de què la resolució d’una base de dades s’ha d’obtenir en base l’element més groller. Per tant, mitjançant un “resampling”, com pot ser la interpolació, no podem millorar la resolució de la base de dades. En el cas anterior, si volem convertir les retícules de 100 m, la resolució serà necessàriament de 100 m. Però si volem convertir-les a una resolució de 10 m, hem de tenir en compte que la resolució és de 100 m encara que la retícula correspongui als 10m.

“In general and from the perspective of product accuracy [...] a spatial database should represent reality, using data which is a raw and unabstracted as possible, rather than the cartographic representation of reality, which is often highly abstracted. If the cartographic view is necessary, it can be generated from the database using appropriate rules. It is much easier to assign

indices of accuracy to raw data than to abstractions and interpretations. (Goodchild 1988:42)

3.3.3. L'ajut dels mètodes de l'estadística i de l'anàlisi espacial

En l'anàlisi de les dades (que consisteix en el processament i interpretació d'un determinat conjunt de observacions), l'aplicació de tècniques estadístiques a les dades numèriques, és el que es coneix com a anàlisi estadística de les dades. En les dades geogràfiques, la component espacial fa que les tècniques pel seu tractament siguin específiques. L'estadística espacial fa referència a aquest àmbit més restringit de l'estadística, que permet el tractament de les dades espacials, basant-nos en les seves distribucions geogràfiques, com són la dispersió i els patrons característics d'un fenomen o procés geogràfic. Aquestes tècniques d'estadística espacial, o també anomenada geoestadística, permeten: la seva comparació, a través de la comparació de la coincidència de les dades corresponents a les observacions recollides dels patrons de dispersió espacial i de l'examen de les relacions i de la seva connectivitat existent entre les diferents entitats o fenòmens geogràfics. ([Walford Nigel, 1995 #489]:346).

A pesar de les crítiques rebudes per la geografia quantitativa, i l'estadística espacial dels anys 70', per donar massa importància a la part descriptiva i massa poca a l'explicació dels processos, en els anys 80' i 90' la possibilitat d'integrar les tècniques d'estadística espacial en els SIG amb la possibilitat de poder visualitzar d'inmediat el mapa resultant, és d'innegable interès per al investigador.

La divisió entre tècniques descriptives i les inferencials o d'anàlisi de patrons, continuen essent vàlides en aquest camp de l'estadística espacial; però tant en un cas com en l'altre es centren principalment en la importància dels aspectes de selecció i localització de les entitats.

És evident que la naturalesa de la informació i la manera de representar-la mitjançant els diferents objectes espacials, requereixen de diferents tècniques estadístiques; a on el concepte bàsic consisteix en considerar que en un lloc determinat en un moment determinat a la distribució geogràfica d'un fenomen, te un lloc fix (únic) que pot identificar-se a través de un sistema de referenciació espacial; a on el sistema de referència implica unes unitats de mesura que a la vegada es basen tant la distancia entre diferents fenòmens, com interna de variació. Son les tècniques estadístiques que

utilitzem en la classificació de dades provinents de sensors remots, que permeten distingir diferents fenòmens en base a la distància existent entre diferents característiques geogràfiques, així com el grau de variació interna d'aquestes característiques. Així, doncs, especialment en el camp d'aquest tipus de dades es fonamental poder comptar amb aquesta tècnica amb llarga tradició però que en els SIG proporcionen nous resultats.

Però les tècniques estadístiques no són suficients per resoldre gran part dels problemes que es presenten durant la generalització. Per exemple, els programes de tractament estadístic a través principalment de la representació de coropletes en molts casos és errònia. La classificació que fan la majoria de programes en base al quantils, segons MacEachren, no és aconsellable aplicar-la per defecte, ja que no es té en compte normalment quina és la distribució de les dades. En el processament de dades en el camp cartogràfic no hi ha la tradició que existeix en el de la teledetecció, que utilitza normalment el tractament de tipus estadístic en les seves dades, el que permet assegurar examinar les dades prèviament a la seva codificació definitiva en un objecte digital.

Intentem que cada classe que pensem representar posteriorment amb un mateix símbol, contingui valors tan similars com sigui possible, i que les diferents classes que simbolitzem de diferent manera siguin el més diferent possible.

"minimization of variation within classes and maximization of variation between classes" MacEachren 1994:50

Una possibilitat que ens permet determinar els diferents punts d'inflexió de la distribució espacial de les dades, és el mètode de Jenks de classificació "optimal statistical data classification", especialment útil en el tractament de gran volum de dades, basat en mesurar la variança dins i entre categories, minimitzar la variació numèrica dins de les categories i maximitzar la variació entre elles. Un tipus de tractament equivalent és el que efectuen els PCA o components principals de l'imatge provinent d'un sensor remot, que permet comprobar el grau de correlació entre els fenòmens que triem en un posterior anàlisi.

El que no té en compte el mètode de Jenks és la proximitat espacial dels valors i els problemes derivats de la necessitat de comparació temporal de múltiples. I no ho té en

compte ja que aquests problemes són deguts a que es un mètode no-espacial de tractament de les dades. Una possibilitat és la utilització dels procediments de clustering que permet agrupar els valors de les dades que són similars i estan pròxims.

Segons MacEachren el mètode que ens permet comparar dos es fer un anàlisi exploratori amb dades que no estiguin classificades

Les possibilitats d'utilitzar les estadístiques espacials no ha de fer oblidar que s'apliquen a dades espacials, ja que la manera com està estructurat un problema pot afectar les estadístiques descriptives resultants. L'avantatge és la de poder resumir la informació espacial a través de la visualització d'histogrames, així com a través de resums numèrics o quantitius. Un conjunt de dades pot ésser resumit de diferent manera: mesures de la tendència central (el valor típic d'una distribució de freqüència), de la mesura de dispersió (tots els números que representen la variabilitat de les dades) i la mesura de la forma o posició relativa (tots els números que permeten descriure la naturalesa o forma d'una distribució de freqüència). La tria del tipus de descripció estadística per un determinat problema depèn del nivell de mesurament: nominal, ordinal o de interval-ratio, a la vegada que varia el procediment de càlcul segons que les dades estiguin agrupades o no.

Així mateix, l'estadística descriptiva espacial també ofereix moltes possibilitats per donar una idea concisa de les característiques d'un conjunt de dades, ja que permet utilitzar les dades originals. I, en el camp del preprocessament de les dades provinents de sensors remots, la possibilitat de visualitzar les dades gràficament faciliten molt seu processament.

3.3.3.1 La variança i el coeficient de variació

La dispersió, la variabilitat i la variança ens donen una mesura del grau de dispersió d'un conjunt de valors al voltant de la mitjana, útil com a estadística descriptiva en un procés inferencial. Però, en realitat, en el cas que aquí es planteja, que és el de poder trobar un mètode que permeti prendre la decisió del sentit que pot tenir integrar o federar dues bases de dades, el concepte del coeficient de variació pot ser el més adequat. Aquest ens permet comparar la quantitat de variabilitat en diferents patrons espacials directament per poder decidir el que té una mes gran variació espacial o temporal. Utilitzar la desviació estàndard o la variança entre llocs o regions no es massa

adequat perquè ofereix valors absoluts, i els valors depenen del tamany amb que han estat calculades (un conjunt de dades amb valors absoluts grans quedaran descrits com a gran desviació standard o varianza, i a l'inrevés). No serveix per comparar la variació espacial entre "dos patrons espacials" diferents i, en canvi, la mesura relativa representada per el coeficient de variació sí que ens ho permet. El coeficient de variació és la desviació estàndard expressada en relació a la magnitud de la mitjana, i normalment ve expressada com a percentatge d'aquesta, el que permet una comparació entre diferents conjunts de dades. Hem de tenir en compte que solament pot ésser utilitzat el ratio-scale i no l'interval d'escala ja que el coeficient de variació és un ratio de la mitjana.

El coeficient de variació com a mesura de variació relativa ens permet resumir de manera quantitativament adequada mapes i altres patrons espacials, i ens permet mesurar i comparar, per tant, les variació en les dades.

Però quan apliquem estadística descriptiva a dades espacials, hem de tenir en compte l'efecte que té la delineació del contorn i el lloc on es troben localitzades. També hem de tenir en compte l'efecte que pot tenir l'alteració de les subdivisions internes (les anomenades "Modified Areal Units Problem", MAUP) en la mateixa àrea d'estudi, així com l'impacte d'utilitzar diferents nivells d'agregació espacial o diferents escales. Aquest seria en realitat un problema relacionat amb l'escala.

L'impacte que té el canvi del contorn extern de l'àrea d'estudi repercuteix en les subàrees internes i, per tant, hem de fer especial atenció a la estadística descriptiva. Podem començar considerant en primer lloc l'impacte en el tamany de l'àrea en xifres absolutes, com són la mitjana i la desviació estandard o la varianza no oblidant, que queden afectades per el canvi de contorn.

El geògraf, quan fa una subdivisió dins d'una àrea, ha de tenir en compte que les estadístiques descriptives resultants variaran en funció de la configuració del contorn extern de àrea, i ens hem de preguntar, per tant, sobre la possible biaix espacial d'una determinada variació interna

En tercer lloc, no hem d'oblidar el problema de l'escala que, en les variables socioeconòmiques es presenta com a diferents nivells d'agregació espacial, i que també repercuteix en estadística espacial, fent que varii l'estadística descriptiva.

Els mètodes exploratoris d'anàlisi espacial són molt adequats per ésser integrats en un sistema d'informació geogràfica, i se'ls hauria de donar un major èmfasi, ja que es centren en la necessitat de desenvolupar detectors de patrons espacials que reuneixin les condicions de: que no siguin dependents de l'escala; que puguin ser fàcilment automatitzats o aplicables en el programari; i que tinguin la possibilitat d'incloure coneixement com el de les persones. Seria de gran importància el trobar mètodes que explorin les bases de dades de manera

“fluid, artistic, and creative [...] and in a way that stimulates the user's imagination and intuitive powers.” (Openshaw 1991:)

Existeix, doncs, la possibilitat de desenvolupar patrons espacials dels fenòmens físics i socials i dels processos que han creat aquestes distribucions. S'estaria passant a considerar els processos probabilístics en lloc dels determinístics, a predir les futures ocurrències dels patrons espacials, que mostrin la variació espacial de la variable considerada

Armstrong, M. P. (1991). “Knowledge classification and organization.” *Map generalization: Making rules for knowledge representation*, B. P. Buttenfield, R. B. McMaster, eds., Longman Scientific & Technical, London, 86-102.

Brassel, K., Weibel, R. (1988). “A review and conceptual framework of automated map generalization.” *International Journal of Geographic Information Systems*(2), 229-244.

Burrough, P. A. (1986). *Principles of Geographical Information Systems for Land Resources Assessment*, Oxford University Press, NY.

Buttenfield, B. (1995). “Object-oriented map generalization: modelling and cartographic considerations.” *GIS and Generalization. GIS Data 1*, J. C. Müller, J. P. Lagrange, R. Weibel, eds., Taylor & Francis, London, 91-105.

Buttenfield, B. P. (1991). “A rule for describing line feature geometry.” *Map generalization: Making rules for knowledge representation*, B. P. Buttenfield, R. B. McMaster, eds., Longman Scientific & Technical, London, 150-171.

Dangermond. (1988). “Survey of computerized Land Information Systems.” , Nieman Vonderohe

Sullivan, University of Wisconsin-Madison.

Dangermond, J. (1990). “A classification of software components commonly used in

GIS." Introductory readings in Geographic Information Systems, Peuquet D. J., M. D. F., eds., Taylor & Francis, London

New York

Philadelphia.

Dutton, G. "Pathways to Sharable Databases." *Auto-Carto-II*, Minneapolis, 157-166.

Flowerdew, R. (1991). "Spatial Data Integration." *Geographical Information Systems*, D. Maguire, M. Goodchild, D. Rhind, eds., Longman, London, 375-387.

Frank, A. U., Mark, D. M. (1991). "Language Issues for GIS." *Geographical Information Systems*, D. J. Maguire, M. F. Goodchild, D. W. Rhind, eds., Longman, London, 147-163.

Gatrell, A. C. (1991). "Concepts of space and geographical data." *Geographical Information Systems*, D. J. Maguire, Goodchild, M. F., and Rhind, D. W., eds., ed., Longman, London, p. 111-117.

Goodchild, M. (1988). "The issue of accuracy in global databases." *Building databases for global science*, H. Mounsey, R. F. Tomlinson, eds., Taylor and Francis, London, NY, Philadelphia, 31-48.

Imhof, E. (1972). *Thematische Kartographie*, Walter de Gruyter, Berlin-New York.

Kelly, M. O. (1994). "Spatial analysis and GIS." *Spatial analysis and Gis*, S. Fotheringham, P. Rogerson, eds., Taylor and Francis, London, Philadelphia, 65-79.

Lacoste, Y. (1977). *La geografía: un arma para la guerra*, Anagrama, Barcelona.

Lagrange, J. P., Ruas, A. "Geographic Information Modelling: GIS and Generalisation." *Sixth International Symposium on Spatial Data Handling*, Edinburgh, 1099-1117.

Langran, G. (1992). *Time in Geographic Information Systems*, Taylor & Francis, London.

MacEachren, A. M. (1994). *Some truth with maps: a primer on symbolization and design*, Association of American Geographers, Washington.

MacEachren, A. M. (1995). *How Maps Work. Representation, Visualization, and Design*, The Guildford Press, New York, London.

Mark. "Conceptual basis for geographic line generalization." *Auto-Carto 9. Ninth International symposium on Computer-Assisted cartography*, Baltimore, 68-77.

Martin, D. (1991). *Geographic Information systems and their socioeconomic*

applications, Routledge, London and New York.

McMaster, R. B., Leitner, H., Sheppard, E. (1997). "GIS-based Environmental Equity and Risk Assessment: Methodological Problems and Prospects." *Cartography and Geographic Information systems*, 24(3), 172-189.

McMaster, R. B., Shea, K. S. (1992). *Generalization in Digital Cartography*, Association of American Geographers, Washington D.C.

Monmonier, M. (1993). *Mapping it Out*, The University of Chicago Press, Chicago.

Muller, J.-C. (1991). "Generalization of spatial databases." Geographical Information Systems, D. Maguire, M. Goodchild, D. Rhind, eds., Longman, London, 457-475.

Muller, J. C. "Challenges Ahead for the Mapping Profession." *Auto-Carto 9*, Baltimore, 675-683.

Muller, J. C. "Mechanisms of geometric abstraction." *5th International Symposium on Spatial Data Handling*, Zürich, 123-133.

Müller, J. C., Lagrange, J. P., Weibel, R. (1995a). "GIS and Generalization." GIS Data 1, I. Masser, F. Salgé, eds., Taylor & Francis, London.

Müller, J. C., Weibel, R., Lagrange, J. P., Salgé, F. (1995b). "Generalization: state of the art and issues." GIS and Generalization, J. C. Müller, J. P. Lagrange, R. Weibel, eds., Taylor & Francis, London, 3-18.

NCGIA. (1990). "Core Curriculum." .

NCGIA, Goodchild, M., Mark, D., Egenhofer, M. J., Kemp, K. K. "Varenius." Proceedings of the Joint European Conference and Exhibition on Geographical Information, Vienna.

Nyerges, T. (1991). "Representing geographical meaning." Map generalization: Making rules for knowledge representation, B. P. Buttenfield, R. B. McMaster, eds., Longman Scientific & Technical, London, 59-85.

Openshaw, S. (1991). "Developing Appropriate spatial analysis methods for gis." Geographical Information Systems: Principles and applications, D. J. Maguire, M. F. Goodchild, D. W. Rhind, eds., Longman Group
John Wiley & Sons, New York, 389-402.

Oxborrow, E. (1989). *Databases and Database Systems*, Chatwell-Bratt Ltd, Bromley, UK.

Peuquet, D. J. (1990). "A conceptual framework and comparison of spatial data

models.” Introductory readings in Geographic Information Systems, D. J. Peuquet, D. F. Marble, eds., Taylor and Francis, London, New York, Philadelphia, 250-285.

Reichenbacher, T. (1996). “Eine Methode für den Wissenserwerb in der Kartographischen Generalisierung durch Interaktionsaufzeichnung und Induktives Lernen,” , Universität Zürich, Zürich.

Robinson, A. H., Sale, R. D., Morrison, J. L., Muehrcke, P. C. (1984). *Elements of Cartography*, John Wiley and Sons, New York.

Robinson, A. H., Sale, R. D., Morrison, J. L., Muehrcke, P. C. (1987). *Elementos de Cartografía*, Omega, Barcelona.

Ruas, A., Lagrange, J.-P. (1995). “Data and knowledge modelling for generalization.” Map Generalization, Müller J. C., Lagrange J.-P., W. R., eds., Taylor & Francis, London Bristol, USA.

Shea, K., McMaster, R. “Cartographic generalization in a digital environment: when and how to generalize.” *Ninth International symposium on Computer-Assisted cartography*, Baltimore.

Spiess, E. (1982). “Kartographie Grundzüge.” Zürich.

Spiess, E. (1990). “Generalisierung in thematischen Karten.” *Kartographisches Generalisieren*(10), 63-70.

Tobler, W. “Resolution, resampling and all that.” *Buiding databases for global science*, First meeting of the IGU global database planning project, Tylney Hall, 129-137.

Unninayar, S. “The global system, observing and monitoring change, data problems, and management and databases.” *Buiding databases for global science*, *First meeting of the IGU global database planning project*, Tylney Hall, 357-377.

Weibel, R. (1991). “Amplified intelligence and rule-based systems.” Map generalization: Making rules for knowledge representation, B. P. Buttenfield, R. B. McMaster, eds., Longman Scientific & Technical, London, 172-186.

Weibel, R. (1995a). “Map generalization in the context of digital systems.” *Cartography and Geographic Information systems*, 22(4), 259-263.

Weibel, R. (1995b). “Three essential buiding blocks for automated generalization.” GIS and Generalization, J. C. Müller, J. P. Lagrange, R. Weibel, eds., Taylor & Francis, London, 56-69.