

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



5. La generalització com a problema: una proposta metodològica

“Today the user is the cartographer. [...] This point cannot be stressed enough”
(Morrison 1994:4)

El canvi que ha suposat la introducció de la tecnologia digital en el procés cartogràfic ha potenciat l'accés, en la pràctica d'aquesta disciplina, a un nou tipus d'usuari, en el que s'ha vingut a denominar la “democratització de la cartografia”.

El pas del tractament de la informació espacial d'un entorn manual a un de digital ha suposat, pels anomenats SIG, ja no un mer canvi tecnològic sinó també teòric i, sobretot, disciplinar. L'usuari, actualment, ja no utilitza mapes analògics sinó que accedeix a bases de dades, i a uns programes que li permeten confeccionar un producte “a la mida”.

Poder automatitzar el procés d'elaboració d'informació és l'objectiu pretès de fons, plantejat de moment a través en el tractament de les dades amb l'ajut del *software*. Les maneres d'aconseguir aquest objectiu de l'automatització han estat diverses i han anat des de l'ajuda a l'usuari en el processament fins a intentar aconseguir unes dades intrínsecament més intel·ligents que facilitin el seu processament.

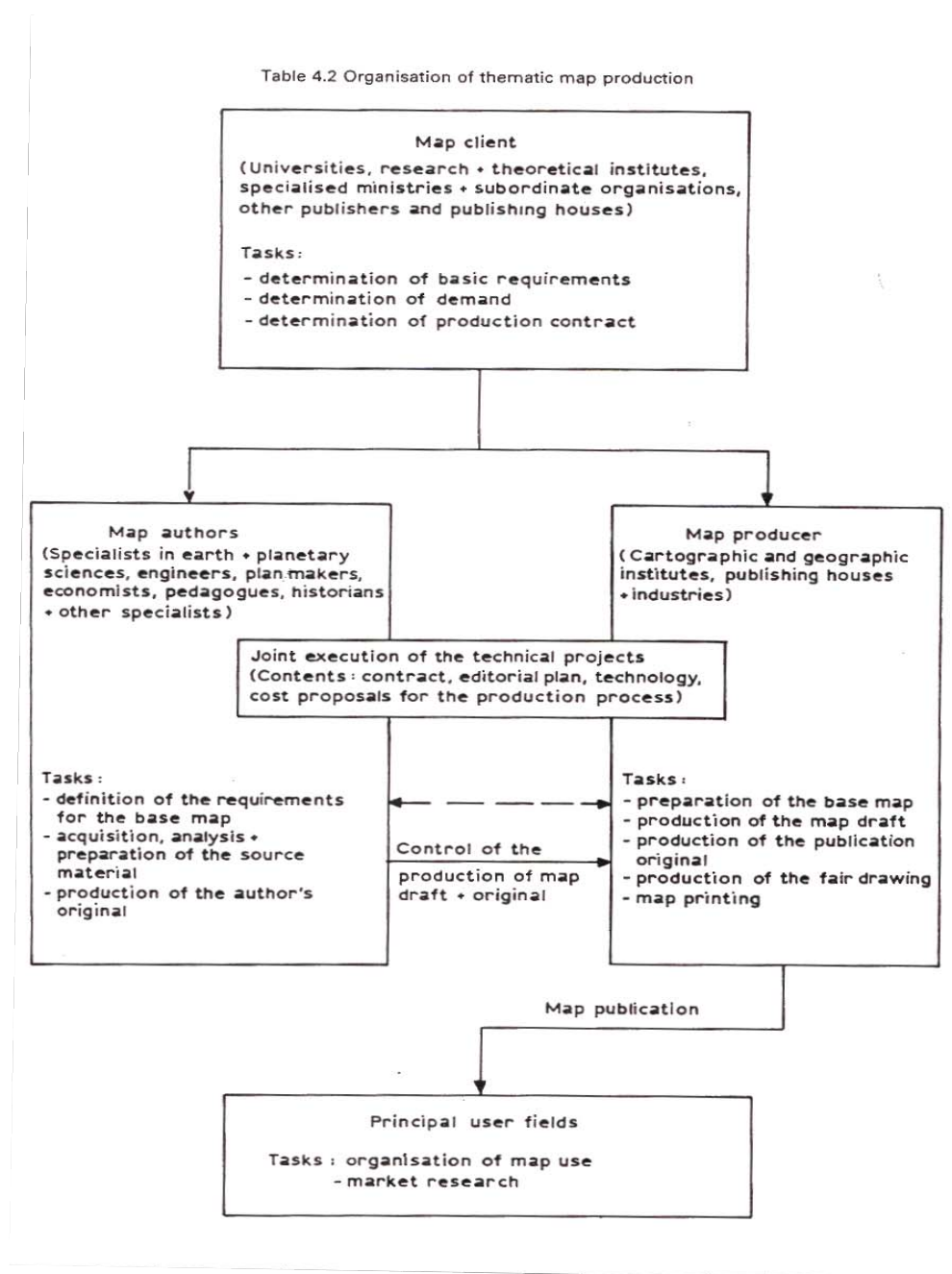
Una via que s'ha mostrat especialment prometedora ha estat la que ofereixen els sistemes basats en xarxes neuronals, que permeten extreure solucions-tipus a problemes concrets, a través de l'aprenentatge de les solucions que aporten els cartògrafs experimentats durant el procés cartogràfic.

La manca de saber cartogràfic s'intenta compensar amb sistemes més intel·ligents, que contenen més coneixements, amb interfícies més amigables i capaces d'acompanyar l'usuari en les diferents opcions cartogràfiques i que li ofereixin biblioteques de processos d'ajut en la resolució de problemes equivalents al que tindria si pogués comptar amb el coneixement d'un cartògraf experimentat.

En un entorn digital, els problemes que se'ns plantegen ja no són de tipus

tècnic sinó de contingut: el de decidir quin tipus de coneixement cartogràfic és necessari. Actualment, la majoria de SIG no ofereixen resistència a tractaments incorrectes; en el intent de reproduir de manera automàtica el procés, desapareix el compromís tradicional existent entre el tipus de mapa que volia el cartògraf, i la informació de la que disposava. Hi ha una falta de comprensió de quines són les limitacions de les dades originals, durant el procés de compilació, degut al desconeixement del procés cartogràfic, ja bé per part de l'usuari o en la compilació (contingut, detall, simbolisme i representació basada en el percebut), per aconseguir una finalitat desitjada. Aquesta manca de coneixement de l'eina de treball es manifesta tant en la manipulació estadística del model de les dades com en la seva representació cartogràfica.

Així ho mostra Ozbeck, quan planteja la tensió existent entre el intent de representar el màxim d'informació per obtenir coneixement sobre un procés geogràfic i les possibilitats reals de les dades, especialment en el del grau de correcció permesa per aconseguir una finalitat determinada.



Font: Lehmann, Ogrisek 1988

Aquest control sobre la decisió de quin és el procés adequat en funció de la finalitat i de les dades de que disposem era més senzill quan el que es manipulava eren mapes, o sigui representacions cartogràfiques de les entitats espacials. Però si el que hem de manipular és el model de les dades, el procés es torna més difícil de controlar i de seguir. En el procés cartogràfic tradicional, la informació cartogràfica podia ser llegida i analitzada visualment, tot i que aquest no sigui un procés “natural” a la persona sinó consistent en un llenguatge tancat i moltes vegades críptic, eliminant

“soroll” a través d’una bona pràctica cartogràfica.

La crítica al model de comunicació cartogràfica dels anys 70 es feu durant el decenni següent en base a la crítica del model de comunicació, ja que es basa en un llenguatge cartogràfic antropocèntric i, per tant, discriminatori. No tothom posseeix els coneixements necessaris per fer una lectura del mapa, i encara que el llenguatge cartogràfic, fundat en la “*Gestalt*” es basa en les capacitats perceptives, no es pot considerar un llenguatge universal.

“To understand the purpose of a modern GIS, the role of the map is an adequate guide. The map product serves as a visual channel of communication, but it must be interpreted inside its frame of reference to impart meaning.” (Chrisman 1987:36).

La introducció de tècniques digitals ha intentat esborrar el component subjectiu del procés, intentant simular un processament neutre de la informació. Aquesta innecessària pretensió de neutralitat en l’elaboració de la informació es manifesta en no voler incloure com a part integrant del procés cartogràfic a la selecció dels fenòmens i processos que s’han d’analitzar. I és dins del procés de generalització on es manifesta més clarament aquesta inhibició pel que fa referència a la selecció, ja que l’eliminació, que seria l’operació contrària a la selecció, en canvi, sí que es considera en el procés cartogràfic tradicional com a subprocés de la generalització. La necessitat de poder automatitzar aquestes dues funcions d’eliminació i selecció de la informació, la tria correcte d’un model de la realitat, ha suposat una de les majors dificultats en el procés digital, que a la vegada ha vist la impossibilitat d’obviar ja que la repercussió d’aquestes funcions en el resultat final és decisiva. La crítica que s’ha fet al treball del cartògraf de no ser neutre respecte a la selecció de la informació, en l’entorn dels SIG no ha minvat ans al contrari, s’ha fet evident que la manca de coneixements cartogràfics per part de l’usuari de quin ha de ser el tractament adequat per visualitzar la informació, són font de greus errors a nivell de processament i de visualització, i per tant de comprensió final de la informació.

El preprocessament representa una ampliació del procés tradicional cartogràfic que possibilita noves maneres de tractament previ de les dades en el model de les dades dirigides a les necessitats d’anàlisi i no a les de comunicació. El tractament previ de les dades ja existia en el món cartogràfic manual (el procés de compilació en podria ser un símil), que no disposava de la possibilitat de manipulació digital prèvia

de les dades. El preprocessament ens permet fer una sèrie d'operacions que en la majoria de casos es realitzaven en el cap del cartògraf en base a l'experiència en el tractament d'un tipus de dades, a través de manipular les entitats del món real simbòlicament o bé mitjançant una sèrie de tècniques una de les quals és la disposició ordenada dels diferents atributs en la matriu de Bertin (Bertin 1977).

El preprocessament es fa principalment en funció de les necessitats d'anàlisi i suposa una manipulació prèvia i a un nivell més abstracte que permeti el "pensament visual" (*visual thinking*), a través de la visualització d'una imatge no simbolitzada.

La situació actual de diferents nivells de tractament de la informació, permet el símil amb la cartografia tradicional, en la diferent intervenció en el procés d'elaboració tradicional del mapa, entre els anomenats geògraf-enginyer- cartògrafs i tècnic-cartògrafs. Aquesta diferenciació ha estat molt evident en la tradició cartogràfica militar, on la jerarquització diferenciava de manera molt nítida, entre els qui dissenyaven els continguts del mapa, en seleccionaven els fenòmens del món real i el tipus de finalitat desitjada de cara a prendre un tipus de decisió, i els que executaven el treball pràctic cartogràfic, formant entre ells compartiments professionals estancs¹.

En l'entorn digital la diferència en la intervenció és molt semblant al procés tradicional, permet l'accés a la manipulació de les dades en dues fases del procés d'elaboració del mapa, el DLM i el DCM, que es podrien caracteritzar per dos tipus d'usuaris. D'una banda, l'usuari que està acostumat a manipular dades espacials i que intervé en el preprocessament de les dades basant-se en el tractament estadístic principalment i en la visualització de les dades que li permet un "pensament visual" MacEachren et al. 1992 i a qui li interessa poder controlar el preprocessament en el model de les dades, però que no coneix, en canvi, les possibilitats del llenguatge cartogràfic de representació espacial. I, per l'altra banda, l'usuari al que l'interessa poder pensar i fer el preprocessament de la informació espacial basant-se en la seva representació cartogràfica.

¹ Això també és evident en la mateixa estructuració de l'ensenyament dels cursos de cartografia en les escoles especialitzades, l'ITC holandès en podria ser un exemple, ja que ens permet veure que els diferents nivells existents entre els cursos de cartografia corresponen a una línia que separa el treball manual de l'intel.lectual, quan en la realitat la pràctica no porta mai a una diferenciació tan exagerada. Per exemple, quan s'ha de solucionar problemes deguts a la congestió d'elements i es necessari desplaçar-los o eliminar-los, una feina en principi complexa i corresponent a una decisió sobre els continguts, realitzada normalment per cartògrafs-tècnics que tenen experiència. Ormeling també evidencia aquesta situació en l'existència dins d'un servei cartogràfic, de diferents nivells jeràrquics, entre els que dissenyen la informació i els que confeccionen el mapa Ormeling 1984.

Són dos processos de manipulació de la informació: un basat en la possibilitat de manipular la informació espacial dins del model de les dades, i l'altre representat per la conveniència en la manipulació de la informació espacial en base a la representació cartogràfica.

El que “dona formes”, el “Umformer” d'Imhof, estaria a cavall entre un i l'altre, i és el que en l'antiga pràctica cartogràfica havien aconseguit certes escoles, el de controlar forma i contingut.

En aquest capítol posem de manifest com en la base de compilació de la informació, tant sigui en el model de les dades com en el posterior pas al model cartogràfic, els objectes espacials han de permetre manipular la informació per la seva anàlisi en primer lloc, però tenint en compte que també ha de possibilitar la seva posterior representació cartogràfica. La tria dels fenòmens que seleccionem del món real, és un dels possibles models de dades, que condiciona la posterior anàlisi; reduir la complexitat de la realitat a través d'un procés d'abstracció, ens permet una millor comprensió dels fenòmens o processos geogràfics però només és de les possibles visions. La reducció d'informació per a una millor comprensió, entre d'altres funcions exposades per diferents autors i recollides en el capítol 2, és part essencial del procés de generalització. També hem vist en el capítol 3 i 4 que les dades de que disposem no tenen sentit a tots els nivells conceptuals, i l'ajut en aquesta tasca constitueix una de les prioritats actuals en el camp de la integració d'informació.

En el camí de l'automatització del procés de generalització una dificultat existent és la manca de protocols de tractament de les dades ja que el processament manual és de tipus holístic i de difícil concreció. El model de solució-d'un-problema permet plantejar un marc de treball on identificar les possibilitats o maneres de fer, que provenen de l'experiència o coneixement de les “idees fetes”, sobre les tasques que s'han de desenvolupar. L'usuari parteix d'un marc conceptual del sistema amb el que treballa, i la solució a un problema, el trobar la seqüència d'operadors per acomplir una tasca, com més semblant sigui als seus conceptes, més fàcilment se'n sortirà.

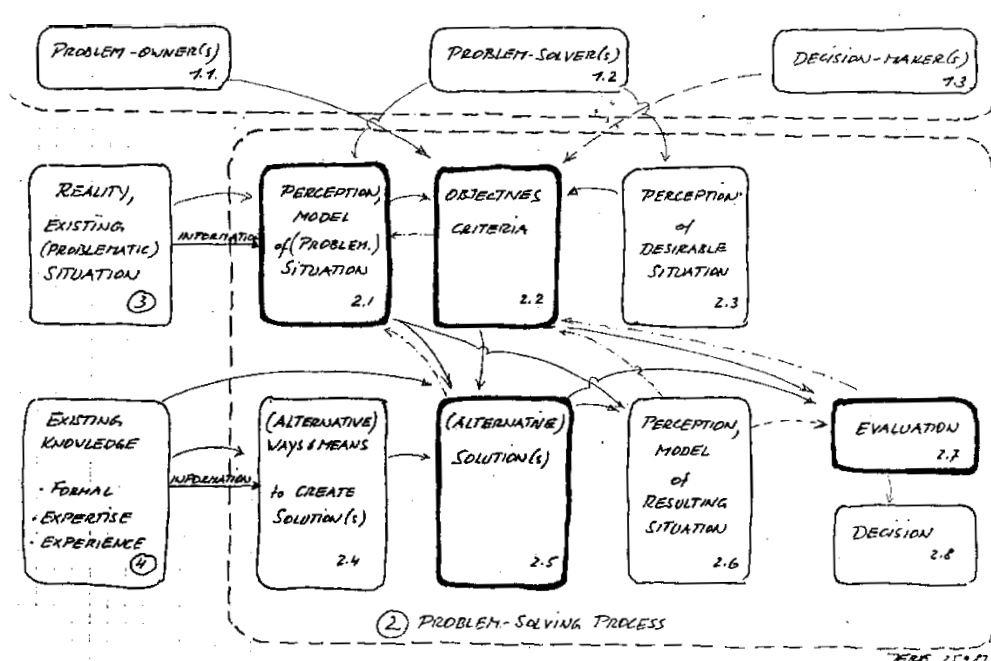
Utilitzem, doncs, com a estructura metodològica la solució d'un cas, procedent del món de la gestió de l'escola de Harvard dels anys 60, que permeti trencar la inèrcia teòrica i la tendència a reproduir el procés manual, i haver d'analitzar l'adequació de

les tasques de la generalització cartogràfica que utilitzem procedents del procés d'elaboració del mapa en un entorn diferent, el digital.

“Modelitzar en darrera”, recórrer el camí invers, permet desprendre's del superflu que pot ser innecessari en l'entorn digital.

5.1 Procés de solució d'un problema

La proposta metodològica no consisteix en el desenvolupament d'una solució a un problema de generalització en un entorn digital, sinó que el que s'intenta és identificar i tipificar el tipus de problema existent en la concreció de certes funcions de generalització en la utilització dels sistemes d'informació, que en el entorn manual eren més fàcilment solucionades.



Font: Procés de solució d'un problema, ITC, De Man, 1989

La dificultat en la generalització de les dades en un entorn digital es caracteritza, per una banda, per la diferent manipulació i conceptualització del que se n'espera del sistema per part d'un usuari que prové de diferents disciplines, i que es troba actualment amb la necessitat de compartir una tecnologia comuna, amb altres

disciplines pertanyents a camps disciplinaris tradicionalment allunyats. És el cas dels SIG, que van conformant la necessitat d'establir una ciència interdisciplinària de la informació espacial, davant de l'evidència de què la complexitat que presenta va més enllà d'un mer problema tecnològic.

S'utilitza l'estructura de la solució d'un problema, per establir un marc de treball corresponent a dues maneres d'afrontar els processos. En el capítol 3 s'ha vist la divergència existent en el que es considera generalització: es confonen els objectius amb el que és el seu processament. I especialment dins del que és el procés de generalització en el procés de compilació, s'ha discutit quines activitats es consideren propiament pertanyents al procés tradicional de generalització: la simplificació, el desplaçament dels objectes espacials cartogràfics, l'exageració, combinació, selecció, classificació, selecció i simbolització.

En un entorn de cartografia manual la dificultat de certs processos és diferent que en l'entorn digital, especialment degut a la diferent intervenció de les persones. Per una altra banda, també s'ha vist en el capítol 2 entre els autors que s'han plantejat la generalització cartogràfica en un entorn digital, Muller assenyala la diferència entre el que ve representat per la generalització geomètrica: simplificació, ampliació i desplaçament, i la generalització conceptual que conté la selecció, la classificació, la tipificació i la simbolització. Aquesta diferenciació prové de l'escola alemanya que sempre ha diferenciat entre la generalització geomètrica i la temàtica, com mostra la figura de Hake referida a les dues principals finalitats cartogràfiques, reproduïda al capítol 1 (pàgina 9).

Aquesta concepció d'un procés de realització d'un mapa participat de dues intencionalitats molt clares basades tradicionalment en l'ús final de la informació cartogràfica té en l'escola alemanya una forta tradició.

Aquests dues maneres d'utilització del mapa són importants tant en el context cartogràfic tradicional com en el digital, però és que, a més, suposen una manera diferent de tractar la informació. Aquest tractament diferent basat en la finalitat del producte cartogràfic comporta una pretensió de manipulació de les dades així mateix diferent, i que es correspon al que hem anomenat el DLM i el DCM. Encara que no és cert que tota la cartografia manual i tradicional es basés en el tractament dels fenòmens del món real en forma de representacions cartogràfiques i que la

manipulació d'aquestes es reduís a la conversió d'objectes primitius geomètrics en funció de la seva finalitat, el que sí és cert és que el pretractament de les dades en la cartografia manual era més difícil de realitzar que no ho és actualment. L'usuari actualment en desconeix el procés tot i tenir les eines, i l'accés de les dades fa que el desconeixement del preprocessament tingui conseqüències més greus en l'entorn digital que no les tenia en el manual. Durant el preprocessament i la compilació decidim la selecció i la classificació i sobretot la tipificació de les entitats inicials, que afecten directament el contingut temàtic del mapa ja que decideixen sobre quin tipus d'informació s'ha de presentar, quines seran les possibilitats posteriors de manipulació i d'anàlisi, i que, per tant, repercuteixen directament sobre el resultat final del mapa. Els criteris sota els que s'han avaluat aquests processos s'han reduït durant molts anys, a la manipulació geomètrica d'objectes considerats de manera aïllada, a on s'ha preponderat el "com" (*How*) per sobre del "quan", el "què" i el "perquè" (*When, What, Why*) proposats per Shea i McMaster (cit Richardson, Muller 1991). Tenir això en compte vol dir deixar de banda la prioritat de la generalització geomètrica, i donar pas una concepció més holística de la generalització, on el procés de disseny de la informació espacial, més encaminada a preveure les necessitats de les tasques d'anàlisi geogràfica, que a la seva representació.

L'estructura de solució d'un problema aporta part de coneixement necessari en la identificació del tipus de dificultats amb les que es troba l'usuari davant del desconeixement del procés de tractament de la informació espacial, en la generalització en l'entorn digital, i que seria aprofitable per poder integrar en les propostes actuals d'ajut per part del sistema a l'usuari. Aquesta aporta de coneixements per part de l'entorn pot ser aprofitada en el que son els Sistemes de Suport a la Decisió (SDSS), i en la instrucció d'un cas en el que es el camp del que s'anomenen CBR o *case-based-reasoning* per la instrucció com a plantilla del que es un cas ben resolt o ben estructurat.

Actualment, els sistemes d'informació geogràfica no tenen eines de decisió que permetin, davant d'un problema espacial -especialment si està mal estructurat- resoldre objectius conflictius, i generar una solució. Segons Dear, aquest fet es deu a que "analysts, rather than decision makers, have selected the dimensions of the problem that are modelled" (citat per Densham, Armstrong 1993:403). Les

possibilitats d'ajut que pot incorporar uns SIG va des dels anomenats SDSS fins als sistemes que formen part del camp de la intel·ligència artificial, i que aporten més ajut a l'usuari. La dificultat de tractament de les dades originals porta moltes vegades a triar variables que no són les adequades i a barrejar nivells de resolució permetin, no tant sols la presa de decisió, sinó una participació del procés de recerca per trobar quina és la solució més adequada.

Per poder realitzar de manera fàcil la presa de decisions, la interfície ha de poder presentar dos espais: l'espai objectiu (que conte els paràmetres i la de la solució i del model analític, gràfics i taules) i l'espai del mapa (que és la representació cartogràfica), i aquests dos espais ha de poder-los interaccionar l'usuari.

De Man fa referència al fet de que la diferencia entre el *problem-finding* o diagnosi i la *solution-finding* en el procés d'estructuració del problema és fonamental i reconegut, i cita els dos tipus ideals de sistemes d'informació: els que estan basats en les dades (en la informació) i els que ho estan en la presa de decisions (Man 1990:333). Els primers proveeixen informació principalment per models no analítics i són la base dels *problem-finding*. Mentre que els basats en la presa de decisions es basen en sistemes analítics, són el que considerem *solution-finding*.

En aquest treball es planteja que la utilització d'aquest marc de treball, permet veure que diferents usuaris "entenen" el procés cartogràfic de diferent manera, per aconseguir l'"homogeneïtzació de l'escala d'anàlisi o d'ús." (Morrison 1994).

Encara pocs *softwares* integren eines de cara a la presa de decisions; un exemple n'és la proposta d'Idrisi feta per Eastman (Eastman et al. 1993:33) que permet aproximar-se a la presa de decisions en una línia bastant semblant a la proposada per l'ITC, també fent evident que és necessari definir criteris i objectius en la solució d'un problema determinat. Segons l'equip de la Clark University per "SIG and decision making" la solució a un problema és vista des del punt de vista més restringit del "decision maker" però amb l'ajut d'un sistema de decisió d'Idrisi. Representa el que es possible plantejar una dècada més tard amb les tècniques participatives d'anàlisi.

"This is a policy decision, made by individuals who perhaps quite nicely fit our traditional concept of decision makers"(Eastman et al. 1993:33)

5.1.1 Identificació del problema

La presentació d'un marc de treball corresponent a la solució d'un problema mostra que la manera com s'han dissenyat els SIG han fet que la seva utilització no hagi estat òptima². Un sistema d'informació ens ha de permetre plantejar una pregunta de manera que la seva resposta sigui assumible pel propi sistema, i si comporta una vaguetat en la seva formulació ha de poder assumir-la. Per tant, un sistema d'informació ha de poder comptar amb unes eines de gestió de la informació que permeti tractar per igual, els elements que el componen com són, en primer lloc, l'entorn; en segon lloc els objectius; i en tercer lloc la manera de dur a terme les tasques.

La generalització és considerada un procés mal estructurat, i com a tal hem de tenir en compte que no podem considerar que ni l'entorn, ni la finalitat ni els procediments siguin clars, com ho serien en els problemes ben estructurats o ben definits. En aquests últims té sentit emfatitzar la busca de solució en base a l'aplicació d'una sèrie de tasques, mentre que en el nostre cas, el dels problemes mal estructurats, és més convenient concentrar-se en la diagnosi i la identificació del problema.

La solució d'un problema és una eina de treball que ja ha estat utilitzada durant dècades en el món de la gestió empresarial però que no és tant habitual en les ciències socials. La seva utilització ens ha de permetre d'apropar-nos a qualsevol tipus de problema, però no esperant una resposta única sinó una resposta probabilística, en especial si el problema que volem plantejar és difícil d'establir, i les dades de les que partim són incertes per naturalesa i si també els objectius són poc definits o fins i tot contradictoris. No podem esperar arribar a una solució única sinó a la més adequada en el seu context, i això seria el més proper al que succeeix en la generalització cartogràfica que és un problema de tipus holístic i difícil de concretar.

L'establiment d'un model de solució del procés de solució d'un problema és complex degut a la interdependència dels elements que participen en el procés, i on és necessari poder comptar amb coneixement extern o amb l'ajut addicional que ens

² En molts casos "GIS ...is not an end in itself. The value of information system arises out of the usefulness of its resultant products. Information is an answer to a question. These questions emerge in the context of problem-solving and often therefore in connection with managerial activities and functions" (Man 1990:325)

informi sobre la naturalesa exacta de la situació existent (el model de la situació existent és normalment la percepció que en té el que s'ha adonat del problema, o *problem-solver*) i de quines són les possibilitats reals d'aconseguir establir el problema (evidenciat per l'estat de la qüestió) en relació als objectius formulats.

Aquest model de solució d'un problema és de tipus funcional i, per tant, els objectius han de poder permetre l'ajut en la formulació dels objectius i fer evident quines són les relacions causa – efecte, però sense deixar de banda el fet de què:

“Problem solving should not attempt to modify reality in such a way that the causes of negative symptoms are altered.” (ITC 1987:12)

També ha de poder ajudar a trobar camins i maneres d'aconseguir els objectius formulats, i ha d'ésser també neutral respecte al tipus de solució permesa. El model ha de poder també comprovar els possibles camins o maneres de solucionar el problema. La solució d'un problema comporta disciplinar-se de manera que es pugui estructurar amb més facilitat a través del seguiment d'una sèrie de regles: 1) concentració en un sol element cada vegada; 2) ésser conscients de que és un procés seqüencial i que, per tant, podem començar per darrera (de la integració de les dades pel que fa a la generalització i la modelització d'aquesta situació); 3) diferenciar entre els processos de pensar divergents (arreglar tants idees com sigui possible) i convergents (anàlisi crítica d'aquestes idees), seguides de la selecció i estructuració de les idees utilitzables; 4) alliberar el subconscient que pot dependre de la intuïció i del coneixement previ; 5) ésser conscients de la jerarquia en els problemes que ens permeti la subdivisió en una sèrie d'elements més petits. El concepte de jerarquia de problemes (del més general al més específic) encaixa bé amb la metodologia d'observar a la vegada els objectius i la situació problemàtica existent. I sense deixar de banda el fet de que hem d'arribar a un punt comú de solució. Per tant, pot ésser que la solució d'un problema suposi la solució de subproblemes i a l'inrevés.

5.1.2 Formulació dels objectius

La generalització ha de permetre un fàcil accés a les dades i, per tant, ha de facilitar determinar l'escala d'integració de les dades; així ho assenyala Weibel quan fa referència a com la generalització del model permet una reducció controlada de les

dades en els dominis espacials temàtic i temporal. Però, apart de reduir l'emmagatzemament i augmentar la rapidesa computacional i una reducció controlada de les dades, ha de permetre l'homogeneïtzació a través de la reducció del grau de correcció i de la resolució dels conjunts de dades.

“Besides data reduction, an important objective of model generalization is the derivation of databases at multiple levels of accuracy and resolution.” (Weibel 1995:57)

El problema que es planteja aquí és ¿de quin tipus és d'impediment que dificulta la concreció d'una escala d'anàlisi a partir d'una base de dades espacials de resolució espacial coneguda i escala mètrica coneguda?

La generalització ens permet derivar un model digital DLM2 d'un altre d'original DLM1, però hem de poder determinar en quin moment s'acompleix l'objectiu, o sigui fins a quin grau de correcció i resolució dos conjunts de dades ens permeten la seva integració³.

“ ...[the questions is]... how much and what kind of model-oriented generalization support is required for the accomplishment of routine tasks in cartography...” (Müller et al. 1995b:5)

Tot i que en l'entorn digital moltes funcions es poden realitzar de manera més senzilla el tractament inicial de les dades és de moment de difícil concreció. El reconeixement de l'estructura i del saber necessari per dur-la a terme no és de moment al nostre abast. Això es deu en gran part a que les funcions que permetrien realitzar-les això, han de poder expressar la complexitat, la distribució i les relacions dels fenòmens cartogràfics, a la vegada que haurien de permetre millorar la selecció i control dels operadors de generalització (Müller et al. 1995b:5)

En el procés d'integració de les entitats seleccionades en el model de les dades, és on decidim si el nivell de detall d'aquestes s'ajusta a la finalitat desitjada. L'escala d'anàlisi s'estableix dins d'un rang de possibilitats, dins del qual les entitats poden ser simbolitzats i codificats de diferent manera, segons siguin les necessitats del seu tractament analític. Si el nivell de detall no coincideix amb el resultat desitjat la

³ Tobler i Moellering demostren al 1972 que la varianza d'una variable espacial ens dona més informació que una resolució simple, amb un alt grau de generalització, que nivells fins de resolució o amb poca generalització.

selecció d'entitats s'ha de modificar, el que pot equivaldre a haver de definir un "espai d'anàlisi" diferent .

La integració de dades esdevé, doncs, una funció prioritària de l'anàlisi en els SIG, i la generalització en l'entorn digital és tant o més necessària com en els sistemes d'informació manuals tradicionals. El procés de compilació de les dades és per tant decisiu i la decisió sobre quina ha de ser l'escala conceptual de tractament de la informació la fa el dissenyador del mapa; és, per tant, una decisió subjectiva que es realitza fora del procés de generalització o del procés cartogràfic

Això ens porta a considerar que:

1) el procés de generalització no el podem basar en l'escala, tal com ve definida tradicionalment com escala mètrica ni en la preservació de la variança, sinó que hem de fer-ho a un nivell superior corresponent a la definició d'un espai més abstracte i associat a una escala de tipus conceptual; hauríem de poder configurar la manipulació de les dades en base al seu significat semàntic i no lògic.

2) el procés de compilació de les dades on es fa l'homogeneïtzació de les dades en base a una hipotètica escala d'anàlisi conceptual decideix la selecció dels fenòmens i quin ha de ser el seu nivell de "sentit" (*meaning*) o nivell d'abstracció

3) hem de poder controlar el procés de generalització que ens porta a la creació d'espais més abstractes, resultants de l'anàlisi espacial i de les possibilitats de tractament inicial de les dades originals del que Muller ha anomenat la "precartografia", on l'absència del concepte d'escala mètrica ens porta a considerar com a relació entre un tipus d'entitat o procés del món real i l'objecte espacial⁴ a l'escala conceptual.

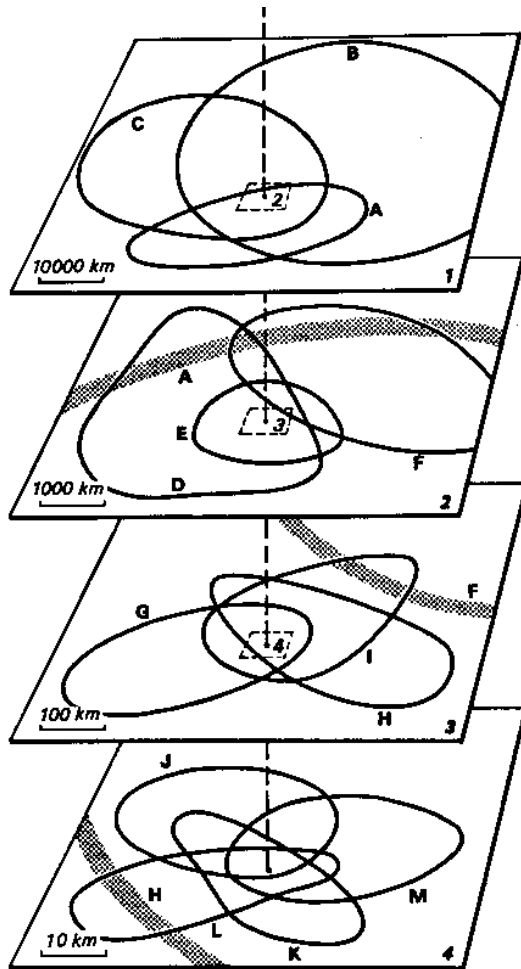
4) aquesta escala es correspon al mapa mental o "visual thinking"

En la pràctica cartogràfica tradicional la integració s'ha resolt definint una escala d'anàlisi i fent la selecció dels fenòmens geogràfics a través de l'adequació d'un conjunt de dades a un nivell de detall i de resolució mètrica, establint un rang

⁴ Això tampoc es soluciona en un entorn d'objecte orientat o sistema d'informació geogràfica virtual, encara que millora, ja que permet que l'usuari sigui conscient del perill i de les conseqüències que té el decidir-se per un tipus d'objecte d'anàlisi o un altre.

d'escala. Hi ha, per tant, una decisió inicial, no relacionada amb la mètrica, que no es basa en la resolució espacial de les dades, que fa referència al nivell de detall dins d'un rang d'escala on té sentit tractar un fenomen o procés geogràfic. I aquesta decisió, dins de l'arquitectura d'un SIG es realitza a nivell conceptual i no a nivell lògic. No tenir això en compte ens porta a la manca de sentit de gran part de correlacions fetes a partir del seu tractament estadístic sobre diferents conjunts de dades, on s'intenta descobrir connexions impossibles, però infinites en les seves possibilitats de combinació. Partir d'un marc teòric suposa acceptar el fet de que partim d'un mapa mental i una estructura cognitiva prèvia, que ens permet definir quin volem que sigui el nivell de tractament de les entitats i de les seves relacions, o sigui quin ha de ser el marc conceptual d'integració, que sorgeix fora del que s'entén com a procés cartogràfic.

El mateix Yves Lacoste fa referència al concepte d'escala i al fet de que no qualsevol fenomen té sentit a qualsevol escala. Això fa que la creença actual d'intentar derivar a partir de la resolució mètrica un nivell d'agregació de les dades o de generalització no ens condueixi necessàriament a "espais possibles" sinó que ens pot conduir a espais d'escala en les que la definició de certs fenòmens no tinguin cap mena de sentit.



(Font: Lacoste 1977)

La dificultat en la integració, prové de la naturalesa mateixa de les dades, i els SIG, actualment, no contenen prou saber per poder fer un tractament “intel·ligent” de les dades originals (1988 Brassel, Weibel 1988).

Una situació desitjable seria la millora de l’entorn, que vindria associada a la facilitat de comunicació entre l’usuari i la màquina:

“An intelligent user interface should be capable of tailoring itself to the needs and idiosyncrasies of individual users, adapting itself to new requirements and new information from the user...users must have control over their own knowledge base profile, in the same way as they have control over their operating system profile.” (Oxborrow 1989:235)

La millora de l'entorn hauria de permetre una comunicació al nivell del llenguatge humà i no del de la màquina. Aquesta millora entorn-usuari és important ja que permet identificar problemes i proposar solucions diferents del procés d'elaboració del mapa en el DLM i el DCM. Aquestes dues parts en que podem dividir el procés de generalització digital corresponen, en primer lloc, a la generalització del model que pretén la reducció semàntica i geomètrica del model de dades original. I, en segon lloc, la generalització cartogràfica dels models dels objectes digitals i als resultats dels processos analítics que permeten aconseguir visualitzacions gràfiques llegibles i entenedores de les dades digitals espacials. (Grünreich 1993).

A part del problema de tipus tècnic hi ha una dificultat en l'accés a dades de qualitat que mantenen els organismes productors de dades, que conserven l'accés discriminat que es tenia a aquestes dades per part de la cartografia militar. En l'actualitat la majoria dels organismes productors de dades han bolcat la informació analògica en paper a productes digitals.

Ara bé, aquest procés de conversió analògica a producte digital, no sempre s'ha realitzat sota un control de qualitat, o de coordinació a nivell estatal, el que és un perill, en especial per un producte que pel fet de que pel seu format digital es més fàcil de ser considerat lliure d'error que el seu original en format analògic.

La falta d'informació sobre les dades o *metadata*, ha comportat errors en el processament posterior, degut al desconeixement dels sistemes de geocodificació de les bases de dades que es van utilitzar en la conversió de format analògic a digital: demanar l'escala pot ser equivalent, en definitiva, a demanar el llinatge de les dades.

Conèixer quines són les necessitats de les dades per la seva la integració en el context de les solucions que pot aportar la generalització en l'elaboració de la informació geogràfica, és actualment una de les qüestions més difícils de resoldre en el context dels sistemes d'informació geogràfica. I això es deu en gran part a la tendència a emular digitalment un procés de tipus holístic, realitzat manualment per cartògrafs, per tant per uns agents diferents dels usuaris dels sistemes d'informació geogràfica. Això fa que gran part de les possibles respostes que ofería la generalització tradicional no s'apliquen en l'entorn digital simplement perquè es desconeixen.

La necessitat en poder integrar dades de diferents fonts, no es un problema de formats per intercanvi de dades, sinó que es planteja a un nivell més complex. Ja hem

vist en el capítol 4 que és dins d'aquest model conceptual a mig camí entre el món real i el model físic de les dades on s'ha de situar la generalització de les dades en base a l'escala.

5.2. Elements de l'entorn

El marc de treball que utilitzem, és un apropament estàtic a un model que s'ha de veure com a dinàmic, on el tractament dels objectes espacials (representacions digitals dels fenòmens geogràfics) a part de tenir en compte la preservació dels processos geogràfics i les seves interrelacions durant el canvi d'escala, ha de poder seqüencialitzar i triar les funcions de generalització i considerar el mapa com un tot.

“Many of the traditional manual generalization processes will be necessary in digital applications, therefore the selection and order of these processes become as crucial as the extent or limit of their application (Stuart-Shea 1991:5)”.

Un SIG està format per un conjunt d'eines, coneixements i persones i és, per tant, aquesta combinació de forces la que conforma les diferents alternatives de treball actualment existent.

5.2.1. Implicació de l'equip

La posició que adopta l'usuari d'un SIG enfront del seu ús és variada. Es combinen a la vegada diferents posicions que corresponen al que té un problema, al que ha de resoldre el problema, i en tercer lloc hi ha el que pren la decisió.

És evident que diferents persones veuen i poden solucionar els problemes de diferent manera, segons una manera diferent de “pensar” i de “comunicar” les dades espacials:

[...] two fundamental changes are occurring simultaneously, not sequentially. Technology and its instrumentation are changing, and also an intellectual paradigmatic shift is occurring”.

Segons Morrison la conseqüència del canvi paradigmàtic és l'aparició d'un

usuari/cartògraf, on el cartògraf es veu separat dels seus productes (Morrison 1994:8.) El canvi de paradigma centrat en l'usuari/cartògraf encara no s'ha produït en la seva totalitat, ja que de moment conviuen el productor/cartògraf i l'usuari/cartògraf. El productor/cartògraf és el que recull les dades, les actualitza a la base de dades per la realització de productes estàndards: és el veritable successor de l'antic cartògraf. El nou usuari/cartògraf té accés a una base de dades espacial i a la funcionalitat d'un GIS, i que com assenyala Morrison, es diferencia de l'anterior:

“[...] such an individual is a true “user” and can create visualizations of his or her own intention, and perform analyses of his or her own fabrication. The electronic technology has empowered this user/cartographer far more than was the case with analogue technology.”(Morrison 1994:5).

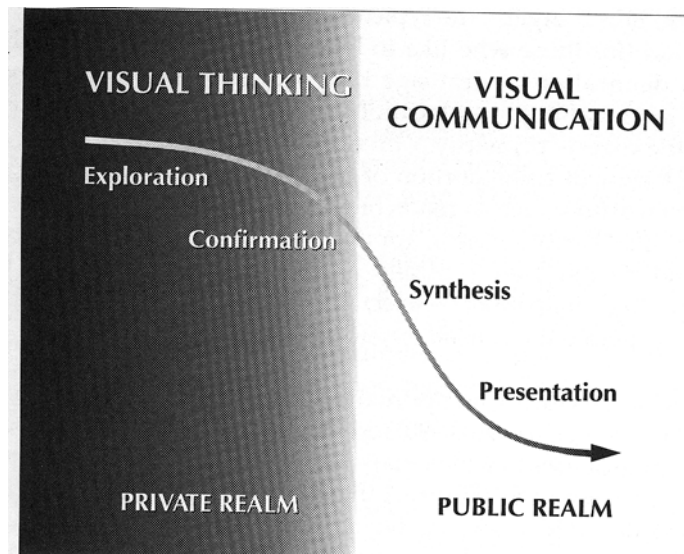
Ara bé, l'usuari d'un sistema d'informació ve representat per l'usuari que utilitza actualment les dades espacials i és un usuari procedent de diferents disciplines i que no està acostumat a la representació cartogràfica de les dades. Un exemple d'això en són els que han utilitzat informació espacial però que, degut a una sèrie d'impediments, disciplinaris de vegades, no estan acostumats a la representació cartogràfica de les dades (per exemple, els economistes o els qui treballen en ciències de la salut). Aquest usuari es diferencia d'un altre tipus d'usuari, que no està acostumat a la gestió del model de les dades, i que prefereix pensar i analitzar la informació espacial a partir de bases de dades cartogràfiques o de mapes. Així com el cartògraf actualment es veu més obligat a involucrar-se en el que és el tractament de les dades, l'usuari científic interdisciplinari necessita coneixements de com poder utilitzar el llenguatge cartogràfic o de comunicació per poder comunicar i representar la informació. Hi ha, per tant, la necessitat de trencar la divisió existent en l'actualitat entre els dos processos de tractament de la informació espacial centrats entorn del DLM i del DCM.

Tenim dos tipus d'usuaris que caracteritzen dos tipus de processament i que tenen una diferent percepció:

Analitza les dades originals i les visualitza i representa	Representa les dades cartogràfiques i les analitza
Les dades son el resultat de l'anàlisi	Les dades son representades pel seu anàlisi
Compila les entitats com objectes espacials pel seu tractament analític i posterior representació	Compila les entitats com els objectes cartogràfics pel seu anàlisi visual
Un producte resultat d'un "propòsit"	Un producte dirigit a "fer evident"

(Font: MacEachren, 1990)

Com mostra MacEachren en la figura següent, hi ha dos tipus de processos que corresponen al que anomena el "visual thinking" i que es regeixen, ja no tant per les regles de comunicació cartogràfica sinó en representacions que elabora l'usuari de tipus més perceptiu, pel control de la seva feina. Per exemple, es pot representar un fenomen que posteriorment simbolitzarem cartogràficament com a àrea i que analitzem com a punt i superfície estadística.



(Font: MacEachren 1994:12)

Aquestes dues fases que diferencia MacEachren es poden identificar en el que representen dins del procés d'elaboració de la informació espacial per un SIG en el model de les dades i en el model cartogràfic.

5.2.2. El procés de solució del problema

És el cor del problema, i és on podem aïllar el problema del seu entorn, i que representa per si mateix un submodel. Els mòduls que tenen una importància a l'hora de prendre una decisió són: la percepció de la situació que ens ha de permetre definir els objectius, els criteris que regeixen aquests objectius i les solucions alternatives que ens han de conduir a una avaluació i decisió sobre el recorregut processual més adequat. Una decisió és una tria entre alternatives basades en diferents hipòtesis, i el que fa el que pren una decisió és l'avaluació entre diferents alternatives basant-se en uns determinats criteris. Un criteri representa un punt de referència que pot ser mesurat i avaluat, i és el que permet fer evident perquè prenem una decisió. Els criteris poden ser de dos tipus: factors i impediments (*constraints*). Un factor el podem definir com un criteri justifica o no l'adequació d'una alternativa específica de l'activitat sota consideració, i que mesurem sobre una escala continua⁵. Un impediment, en canvi, ens obliga a limitar les alternatives que tenim en consideració.

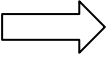
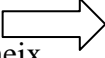
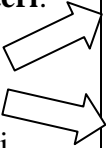
El problema de fons que se'ns planteja en el procés de generalització, i en el cas específic d'integració de la informació, és la dificultat, no tant sols de com generalitzar sinó en saber en quin moment s'ha acabat una tasca, fins a quin punt hem de generalitzar, en quin moment hem de parar l'aplicació d'un algoritme determinat o d'una funció determinada per que ja s'ha aconseguit una escala adequada. Així, doncs, el que resulta de difícil concreció en els sistemes actuals és la concreció dels objectius, de l'escala final d'anàlisi, en absència molt sovint d'una escala de partida de les dades originals. La dificultat en definir les característiques de l'escala final, seria l'equivalent al QUANT hem de generalitzar, que en la cartografia tradicional, es resol a través de la conversió dimensional dels objectes espacials per la seva posterior integració.

Encara que en l'entorn digital s'assenyala el fet que la cartografia tradicional no és un procés que s'hagi de mimetitzar, actualment han passat prou anys com per a que els nous usuaris desconeixin l'antic procés de generalització i especialment de quines són les normes de simbolització i de combinació de les dades cartogràfiques,

⁵ Hem de veure els factors com *variables de les decisions* (variables decisions) en la literatura de programació matemàtica i *variables estructurals* (structural variables) en la literatura de programació (linear goal programming).

en el que s'anomenava la cartografia dinàmica. Per tant, com diu Weibel, moltes vegades degut al desconeixement de la situació anterior s'està redescobrint la roda. I una de les possible solucions que podem aportar actualment és el coneixement cartogràfic corresponent a l'antic procés de generalització i, en especial, al tractament que es donava a les dades originals per l'escala de compilació dels documents originals.

L'antic procés de compilació suposava un preprocessament de les dades originals, a través del seu processament estadístic o "visual" a través de les representacions cartogràfiques ja existents en mapes.

<p>Decisió:</p>  <p>es una tria entre diferents alternatives</p>	<p>Alternativa:</p>  <p>és defineix segons determinats criteris</p>	<p>Un criteri:</p>  <p>factors i impediments</p>	<ul style="list-style-type: none"> ▪ Factors: criteri que confirma l'adequació d'una alternativa ▪ Impediment: és un criteri que nega l'adequació d'una alternativa sota consideració
--	---	--	---

<p>Percepció de la situació problemàtica:</p> <p><i>el procés de compilació de la informació del mapa és dependent de l'escala i del propòsit</i></p> <p>factors necessaris per la solució:</p> <ul style="list-style-type: none"> ▪ establir mecanismes per fer el pas entre nivells d'abstracció ▪ El tipus d'utilització, i el grau de correcció de les dades, de precisió i de resolució de les dades, les hem de triar durant el procés de compilació de la informació <p>identificar les causes dels símptomes negatius:</p> <ul style="list-style-type: none"> ▪ sistemes descentralitzats <p>definir les limitacions:</p> <ul style="list-style-type: none"> ▪ el nivell òptim de detall espacial i el mètode de georeferenciació son funció de les necessitats d'anàlisi <p>definir les oportunitats de canvi:</p> <ul style="list-style-type: none"> ▪ un bon disseny de la base de dades espacial (Calkins 1982) Marble 1988 	<p>Objectius:</p> <p><i>integració de les dades</i></p> <p>En conflicte:</p> <p>objectius necessaris (hard)</p> <ul style="list-style-type: none"> ▪ Conversió de diferents conjunts de dades a través de les seves resolucions o patrons espectrals (nivell lògic): tamany òptim de les unitats i nivell de detall <p>objectius addicionals (soft)</p> <ul style="list-style-type: none"> ▪ Els conjunts de dades de recursos naturals que tenen un alt grau de variació espacial i un baix "decay rate" son més apropiats com dades bàsiques. ▪ Dels conjunts de dades els que son dades bàsiques o primàries tenen una influència mes gran sobre la metodologia d'integració que els relacionats o derivats. (importància del mapa base o dades primàries) 	<p>Percepció de la situació desitjable:</p> <p><i>recollida de dades amb sistemes de geocodificació totalment compatibles</i></p> <ul style="list-style-type: none"> ▪ Sistemes d'informació que serveixen a un grup heterogeni d'usuaris haurien de permetre la fàcil conversió de la informació ▪ Quan integrem conjunts de dades espacials que difereixen en escala o detall s'hauria de poder avisar als usuaris de les limitacions de les dades
---	---	---

<p>Alternatives:</p> <p><i>repensar la naturalesa de les dades segons siguin orientades a l':</i></p> <p>anàlisi</p> <ul style="list-style-type: none"> ▪ repensar els nivells de mesura de les dades i de representació a través dels seus geomètrics primitius <p>síntesi</p> <ul style="list-style-type: none"> ▪ "view integration" \cong virtual gis 	<p>Solucions alternatives:</p> <p><i>recuperar coneixements a els SIG:</i></p> <ul style="list-style-type: none"> ▪ revisió de l'antic procés de compilació cartogràfic i de l'escola quantitativa 	<p>Percepció de la situació resultant:</p> <p><i>prioritzar una visió orientada al model i a l'anàlisi</i></p>
---	--	---

5.2.2.1. Modelització de la situació problemàtica

Necessitem poder generalitzar per poder abstraure i reduir el món real, per poder-lo entendre millor, però en aquest procés de reducció de la realitat a través d'un dels possibles models ens trobem amb que els fenòmens o processos no tenen el mateix sentit a diferents nivells d'abstracció. Els fenòmens o processos del món que seleccionem en el model de les dades, són dependents de l'escala i varien en funció del nivell d'anàlisi desitjat; en el nou entorn la concreció tant de l'escala de partida com la d'arribada són de difícil concreció en l'absència del mapa tradicional imprès sobre paper i representat per les bases de dades.

L'enfocament tradicional de la generalització ha estat principalment orientat cap al camp de la topografia i la cartografia o model cartogràfic on la importància mètrica és essencial. En la cartografia digital s'ha prioritzat també aquesta visió mètrica del que es considera la generalització en "superfície" o representació cartogràfica. I això s'ha fet per sobre de la generalització del model de les dades. La generalització cartogràfica requereix tant de la "informació interna" d'un objecte espacial, de les seves característiques espacials, semàntiques i temporals, com de la "informació externa" que fa referència a les relacions entre objectes i la seva importància contextual. La resolució dels conflictes necessita dels dos tipus d'informació per poder resoldre el problema.

Però en el moment actual aquest concepte restrictiu de la informació geogràfica, ha fet veure que per l'usuari d'un SIG es insuficient ja que es troba amb la dificultat de saber quin és el tractament que ha d'efectuar a les dades originals en absència de les representacions cartogràfiques o escala de referència. La dificultat en poder definir quin tipus de procés hem d'utilitzar en el tractament de les dades, es deu en primer lloc a la dificultat de reconèixer l'estructura mateixa de les dades:

“[...] the user interacts with the geo-objects without knowledge of their internal representation” (Müller et al. 1995a:6)

I això té conseqüències sobre els objectes espacials i, per tant, sobre l'anàlisi i la seva representació; aquesta falta d'atenció, que ja es donava en la cartografia tradicional, fa que en la cartografia digital, com assenyala Muller:

“The way the data model is organized and can be generalized is likely to influence the performance of cartographic generalization” (Müller et al. 1995a:6)

La dificultat pel nou usuari d'estructurar un procés de visualització i representació de la informació espacial a través del tractament de les dades originals en absència d'un preprocessament clar, ja es donava en el procés cartogràfic tradicional. En la cartografia manual, la majoria de vegades moltes representacions s'han fet amb inèrcia, sense el replanteig inicial del tipus de processament o de quines són les necessitats i de quins serien els objectes espacials i nivells de mesura més adequats per l'anàlisi o una millor representació.

Hem de considerar la importància de la seqüencialització de les funcions, així com el control de l'error durant el procés de generalització, que ens permeti decidir millor entre diferents tipus d'operadors. Fins el moment actual el “cóm” fer la reducció d'escala de les dades, la tria de l'algoritme més adequat ha prevalgut per sobre de definir " quin ordre seqüencial i quina quantitat" de generalització s'ha d'aplicar, el que equival a poder definir quina és l'escala adequada d'anàlisi. Aquest és un problema amb el que es troba l'usuari una vegada ha seleccionat un model de la realitat: la dificultat existent en poder determinar quin és el nivell òptim que permet una major comprensió d'un fenomen (Tobler 1988) .

El procés de solució del problema ens ha de poder identificar quin tipus de problema dificulta actualment:

- entendre la importància de l'escala conceptual en el procés de compilació i, per tant, en el preprocessament de les dades

- conèixer la resolució pot equivaler a una escala mètrica i a un nivell de detall, però que l'escala mètrica tradicional portava associada supòsits conceptuals

El problema es pot resumir en: ¿quina és la relació entre resolució i escala? ¿es possible deduir una relació entre resolució i nivell de detall?. Per Joao “to create a level of detail appropriate for the scale of display or analysis” seria una de les raons per generalitzar (Joao 1991). També Goodchild es refereix al nivell de detall i l'escala òptima o possible ([Goodchild, 1997 #545]:2)

Aquesta percepció de la situació problemàtica, és feta principalment des del punt de vista del *problem-solver*, per tant ha d'ajudar en la “formulació dels objectius” per una banda, i en la “trobadura de solucions” per l'altre. En l'actualitat, el problema de la generalització és més evident que en els inicis dels sistemes informàtics degut a la quantitat massiva d'informació disponible. A diferència del que succeïa anteriorment les dades que són al nostre abast són estranyes ja que no han estat recollides en funció de les necessitats geogràfiques tradicionals o per una finalitat determinada, sinó que, en moltes ocasions, són dades pensades per ser utilitzades per diferents disciplines.

Durant els darrers 15 anys la manipulació de dades espacials ha passat per diverses etapes, des de la preocupació en la identificació de quines eren les consultes adequades per determinades aplicacions en les bases de dades espacials; per decidir ¿quins eren els models adequats i quines estructures de dades eren les més adequades?. Un període posterior ve representat per la preocupació en com fer per mantenir aquestes bases de dades cartogràfiques i quin control de qualitat utilitzar en la generació de nova informació a través de la integració de dades ja existents, el que significa oferir un control de qualitat sobre aquest processament o també anomenat “*metalineage*”. Als anys 90 sembla que l'èmfasi se centra en determinar els estàndards adequats per les dades, i en la seva interoperabilitat.

“if we assume that spatial access and search methodologies among multiscale and multisource data bases will be established activity relating the operability among the data is one of the primary issues yet to be successfully resolved.” (Richardson 1996a)

I en aquest sentit Richardson cita la prioritat que s'ha de donar a les maneres en poder resoldre el tipus de processament de les dades espacials de manera més senzilla i sintètica.

Joao també fa referència al fet de que en el moment actual es dona un tractament de generalització igual, tant si els diferents tipus de dades pertanyen a les ciències de la terra o socioeconòmiques, per part dels operadors de generalització (Joao 1991:495). La diferència en l'aplicació en l'entorn digital vindria només en funció del diferent de tipus de modelització de l'espai, raster o vectorial, cosa que no era així en les regles de la cartografia tradicional. La generalització d'una línia que representa una carretera i la que representa un riu, que no corresponen al mateix tipus d'entitat en el món real, no s'han de generalitzar de la mateixa manera i requereixen, per tant, diferents operadors. El tractar per igual dos fenòmens representats per un mateix element geomètric, posa en evidència que estem confonent el fenomen del món real amb l'objecte geomètric espacial que el representa en la base de dades.

"A cartographer when performing generalization needs to make simultaneous decisions about groups features (local level) and about the layout of the map as a whole (global level) (Mackness and Scott, 1988)." (Joao 1991:496)

El cartògraf quan manipula la informació resol i integra dades pertanyents a diferents escales seleccionant, simplificant i exagerant de manera que objectes que no serien vistos a una escala determinada apareixen prenent una importància fora del seu context. Aquesta reconversió de dades provinents de diferents escales a l'escala d'anàlisi, es fa a través d'una sèrie de mecanismes que tenen unes regles estrictes, en el llenguatge cartogràfic.

En les bases de dades no tenim la possibilitat de manipular la informació a partir d'objectes rígidament preestablerts, i la llibertat de poder treballar amb les dades originals és a la vegada, posar al descobert el problema existent de les dades a considerat en absència del mapa en suport paper. El marc del mapa és més que una solució de presentació gràfica, ja que permet acotar una zona de l'espai a partir de la que es seleccionen certs fenòmens, i és un dels passos importants en el procés de compilació d'un mapa.

És ja evident que per poder comprendre el món que ens envolta hem de poder abstraure el superflu del que ens interessa, i que el procés de generalització no es pot recolzar en el concepte d'escala mètrica com havia estat fins a l'actualitat. Però no podem fer-ho amb un tractament de les dades a baix-nivell o nivell lògic sinó que hem de poder fer-ho a un nivell mes alt, de tipus conceptual.

“ Resolving heterogeneous data within the image domain requires generalization, but it must be coupled with semantic integration [...] When different scales or image resolutions are introduced into the integration scenario, spatial and semantic integration must be addressed in the context of scale integration, which involves both statistical and cartographic generalization. These processes are necessary to adjust the different levels of spatial and thematic content” (Richardson 1996b:1).

Per Brassel el canvi principal en un entorn digital, és degut del canvi en els objectius de la finalitat del producte cartogràfic, que fa que aquest ja no sigui un “mapa compromís” sinó diversos mapes o representacions que responen a preguntes múltiples i concretes. La modelització de la informació en un sistema d'informació fa necessari que podem automatitzar, i per tant poder fer explícita, quina son les funcions que han de participar en el nou procés. En la comparació que presenta Brassel entre el procés de generalització manual i el procés en un entorn digital, mostra que en l'antic procés de generalització el cartògraf tria de la realitat els fenòmens per un procés d'abstracció, de quines seran les estructures que les han de representar en el mapa final, i a la vegada imagina les diferents possibilitats de com es poden representar, el que du implícit el procés de reconeixement de l'estructura.

“Ohne eine analoge structureerkennung und modellierung werden automatische Generalisierungsverfahren auch nicht auskommen, wenn sie komplexe Aufgaben sachgerecht erfüllen wollen” (Brassel 1990:40).

És necessari, per tant, poder determinar d'entrada quins són els objectes i les estructures en la base de dades i en els mapes de partida.⁶ A l'entorn digital el procés de tractament de la informació espacial és conceptualment diferent que en l'entorn manual: així mateix ho percep l'usuari, que no reconeix fàcilment quins són els

⁶ Es sind also vorerst Verfahren zu entwickeln, die relevante Objekte und Strukturen in Datenspeichern und auf Ausgangskarten erkennen. (Brassel 1990:40)

processos i funcions que ha de utilitzar.

En el procés de representació cartogràfica traslladem la informació geogràfica a un sistema de símbols cartogràfics, on l'essència del procés de disseny cartogràfic reposa en el reconeixement de la naturalesa de la informació que volem representar en el mapa a través d'utilitzar les propietats perceptuals, que en principi haurien de ser utilitzat per qualsevol persona, ja que es basen en teories perceptives de la *Gestalt*. Aquestes propietats perceptives suposen en si mateixes una generalització de les variables visuals, i corresponen a les propietats associatives, selectives, d'ordre i quantitatives. La correspondència entre el tipus de dades que volem representar i el tipus de símbol que utilitzem fa que distingim tres tipus d'informació: informació qualitativa, d'ordre i quantitativa.

El disseny d'un símbol ha de correspondre el màxim possible a través de atorgació d'una propietat perceptiva a la naturalesa de la informació que representem en el mapa. En el quadre següent mostrem el tipus de informació i la variable visual que utilitzem per representar-la.

<i>Naturalesa de la informació:</i>	Propietats perceptuals de representació
▪ <i>Quantitativa</i>	▪ Quantitativa
▪ <i>Ordre</i>	▪ Ordre
▪ <i>Qualitativa</i>	▪ Associativa i (+/- selectiva)

En la literatura alemanya està ben establert el fet que, en funció de la pregunta que ens plantejem, hi involucrem una o més variables, i en funció de la pregunta que plantejem en el mapa podem diferenciar entre: mapes inventari que corresponen a la resposta a una pregunta basada en una classe de una variable temàtica; mapes per l'ensenyament que permeten establir relacions dins d'un grup de classes d'un component temàtic; i mapes per a la investigació o treball que es refereixen al conjunt de classes d'una component temàtica. Aquest és l'esquema en que es basa MacEachren quan es refereix a la diferents maneres d'utilització del mapa, que resumeix com a "visual thinking" i "visual communication".

Com indica Hake, en el model de generalització cartogràfica on la representació per la comunicació es l'important és fàcil de confondre l'objecte espacial

que hem per “codificar” la informació o objecte espacial i el que utilitzem per representar la informació ja que en aquest model el processament de la informació es suposa que es fa a partir de la lectura. En un entorn digital la necessitat de no confondre l’objecte espacial amb l’objecte cartogràfic, es difícil de recordar i ha estat obviat, moltes vegades en el tractament de les dades. Com assenyala Mark, no és la representació de les dades, el que hem de generalitzar en un entorn digital, sinó el seu model. Els SIG permeten un tractament de les dades que és menys "visible" o més difícil de seguir visualment però que per l'absència d'una escala dels objectes de la base de dades, fa possible que la manipulació d'aquests objectes espacials determinin uns espais de representació que siguin “espais conceptuals impossibles”.

Descriure tots els factors que es necessiten per trobar una solució al problema: *dissenyar el sistema d'informació en funció del tipus d'utilització, un sistema de geocodificació adequat per un tipus d'anàlisi.*

El tipus d'utilització fa necessari que decidim sobre quin ha de ser el grau de correcció, de precisió i de resolució de les dades, i és necessari que això ho fem durant el procés de compilació de la informació.

Nyerges cita la diferenciació que fa Martin de bases de dades basades en el tipus d'aplicació, i les que ho fan en base al tipus de "subject" també anomenades "multipurpose geographic database" basades en la idea de una base de dades de tipus "multipurpose cadastre", més general.

El tipus d'utilització que farem d'un sistema el basem en un disseny específic enfocat a desenvolupar diferents activitats: recerca, inventaris i gestió i/o avaluació. És en base a la concreció de quina ha de ser la finalitat que es pot estructurar el disseny de la informació, que ha d'ajustar-se el màxim possible a la realitat. El disseny d'un sistema d'informació està fet en funció de les necessitats específiques d'una feina, que no té perquè coincidir amb un sistema d'informació més general.

Els sistemes d'informació geogràfica s'han de poder desenvolupar tenint especialment en compte els diferents components: base de dades conceptual, així com els processos per la recollida, emmagatzemament de cara a l'anàlisi i representació de les dades. L'absència d'una necessitat de la representació tradicional impresa del mapa, fa que es perdi el marc de referència del anomenat "enquadrament" del que es

anomenat "la cara del mapa". L'espai que definim d'enquadrament en el mapa i per representar un espai de la realitat, deixa de tenir importància com a relació mètrica en els SIG, però no deixa de ser important, en canvi, per expressar el nivell de detall conceptual al que volem analitzar certs fenòmens o processos.

En el moment actual disposem d'eines que ens permeten manipular les dades però ens és difícil de poder explicitar el procés cartogràfic de generalització. Quan desenvolupem un model, aquest determina les necessitats de les dades i el disseny del contingut del sistema d'informació.

Els inventaris generals que poden ser utilitzats per un conjunt heterogeni d'usuaris per analitzar situacions diferents han de ser flexibles i adaptables, ja que han de poder integrar informació "informal" així com permetre la seva consulta a diferents nivells de detall.

Identificar les causes dels símptomes negatius que s'han de millorar: sistemes descentralitzats

Durant molt de temps, fins i tot dins d'un mateix organisme, s'han desenvolupat dades diferents de manera descentralitzada, però aquesta política s'ha demostrat molt costosa a llarg termini, en especial pel que suposa el manteniment de bases de dades redundants.

Actualment, preval la idea de que cada organisme organitzi el seu propi sistema i la recollida de dades de manera descentralitzada, tot i que seria desitjable que aquesta descentralització fos coordinada, ja que pel fet de no ser-ho es recullen les mateixes dades amb diferents criteris i quasi mai per augmentar-ne la qualitat.

La recollida de dades i el sistema de geocodificació és clau en la posada en marxa i en les possibilitats d'anàlisi posteriors. Anomenem referenciació geogràfica o geocodificació a la identificació d'un fenomen o procés espacial mitjançant un identificador espacial. La tria del mètode més apropiat de geocodificació depèn de les necessitats de la informació segons sigui la finalitat d'anàlisi, tenint especialment en compte l'autocorrelació espacial i la densitat espectral de l'anàlisi.

Podem fer la geocodificació de la informació espacial de manera directa i indirecta, o sigui, donant una referència espacial de manera directa o fent-ho a través d'un altre objecte. La geocodificació actualment es fa separant la informació corresponent a la representació gràfica d'una banda i la que correspon als atributs d'una altra. Els sistema que gestiona aquesta informació no té perquè ser del mateix tipus. Anomenem geocodificació directa a la identificació que fa correspondre els elements geomètrics primitius d'àrees, línies o de punts per la delineació dels diferents fenòmens. Anomenem geocodificació indirecta a la referenciació espacial a través d'una referència espacial a un altre objecte. I tant si són directes com indirectes a les referències geogràfiques les anomenem unitats bàsiques espacials. Quan triem un mètode de geocodificació hem de tenir en compte, entre d'altres característiques espacials, el tipus de variació espacial. Les unitats espacials les podem definir a través de la delimitació dels seus contorns o prenent-ne el centre de gravetat. La variació espacial d'aquestes unitats està estretament lligada al nivell d'agregació. I encara que puguem fer servir un mètode de referenciació geogràfica per diferents nivells de detall espacial, normalment necessitarem diferents mètodes de referenciació geogràfica per conjunts de dades diversos.

El cartògraf, quan generalitza, utilitza saber a diferents nivells local i global (Joao 1991:49) per descriure quines són les necessitats de tractament dels nous objectes en el procés de generalització durant la compilació de la informació (Ruas, Lagrange 1995:76) .

El cartògraf, en el procés de compilació de la informació, en el pas de derivació de l'esquema inicial de les dades Ruas, Lagrange 1995:83) a un nou nivell de percepció, redefineix els objectes més significatius intentant variar-ne al mínim la seva forma geomètrica. El nou tipus d'objecte implica unes possibilitats d'anàlisi i dels operadors de generalització que puguem aplicar a nivell d'esquema; són els de classificació, agregació i associació, i es basen especialment en les relacions i propietats semàntiques entre els objectes.

"When digital data are used, features and even information describing them tend to be regarded independently. This is due to the fact that only a sequence of points and an identification code are needed to represent features graphically...Therefore it is essentially to preserve in the best possible way geometric properties, and spatial and semantic

relations in the process of generalization while respecting graphic limitations which depend on symbolization and the new spatial resolution." (Ruas, Lagrange 1995:76)

Les necessitats de processament de la informació espacial, en base a l'escala i propòsit del mapa, com a principals determinants del procés de generalització, són vàlids tant en l'antic procés com en el nou entorn.

En l'escola suïssa, el propòsit del mapa està ben establert, com ho estaven les preguntes que el cartògraf feia al client, i que el conduïa a la realització d'un mapa específic. Els diferents nivells de preguntes corresponen a consultes elementals, que es centren en una única classe o nivell de la variable o component temàtica (¿quin tipus de sòl hi ha en un lloc?), la segona consulta es fa en base en un grup de classes d'una component (¿quin tipus de sòls son coberts per un aiguat?) i finalment tindriem consultes de tipus integral (basades en la comparació de dos grups diferents de components de les diferents classes). Segons el tipus de consulta podem tipificar els tipus de mapes resultants en: inventaris (no es necessari veure tota la informació d'un cop, son un inventari gràfic), mapes educatius o divulgatius, i mapes de recerca o de treball (que permeten descobrir noves estructures i nivells superiors dels conceptes). Això es important ja que ens conforma el contingut que tindran els diferents tipus de mapa.

La generalització no és motivada només per la reducció en l'escala de representació, que vindria representada per la necessitat de la comunicació, sinó que consisteix en una transformació espacial a partir de quatre necessitats, principalment:

1) Necessitats econòmiques. La recollida de dades es fa amb uns procediments que venen influenciats per unes restriccions econòmiques i tecnològiques. Les bases de dades original, per tant, representa ja una generalització del món real (per Armstrong, una abstracció). Això es fa mitjançant el mostratge discret a partir de mapes digitalitzats o dades agregades en unitats censal. La simplificació i selecció geomètrica o conceptual ens permet reduir el volum de informació que hem anomenat model primari del món real.

2) Necessitats dirigides a aconseguir unes dades "robustes" a través de la correcció dels errors instrumentals i humans procedents de les fonts d'informació com poden ésser els procedents d'una classificació errònia o de la creença que la precisió

en les mesures o major recollida de dades pot fer baixar els errors en la seva interpretació. Per tant, veiem que la generalització ens pot ajudar a filtrar errors que es realitzen al llarg de tot el procés: una base generalitzada és més robusta que una que no ho és.

3) Necessitats variades. Ja hem vist que la generalització és necessària per una base de dades "robusta" però veiem que el filtrat es realitza principalment en funció de les necessitats dels usuaris i de l'escala. Per tant, el que fem és fer un filtrat d'un model original o DLM a un altre model de menys resolució DLM (Grünreich 1985).

4) Necessitats de visualització i comunicació. Aquestes es centren al voltant de la definició de les restriccions en la càrrega de la relació representada per l'espai ocupat pel blanc/negre i l'altre més centrat en els aspectes de comunicació. Bertin proposa una densitat màxima a partir de la qual el mapa ja és il·legible.

Identificar les causes dels símptomes negatius que s'han de millorar:

recollida de dades amb sistemes de geocodificació totalment compatibles

La dificultat en el reconeixement de l'estructura interna de les dades dificulta des del seu inici la possibilitat de compartir les dades i la seva consegüent integració. Si, com és el parer actual, les dades tendeixen a anar cap al que s'anomena com *Multipurpose Geographic Database* o bases de dades normalment a l'escala més gran, és doncs evident que la base de dades no està pensada de cara a la seva aplicació sinó que la derivació a partir d'aquesta ha de ser feta per l'usuari.

Nyerges presenta una metodologia d'integració de la informació que defineix com:

"Information integration is defined in this context as the bringing together of information parts into a working whole; controlling redundancy where appropriate. [...] This bringing together does not mean that the whole is one "physical" whole, however it does mean that the parts cooperate." (Nyerges 1989a:27)

En el procés de compilació de les dades espacials, intervenen de manera decisiva tres elements inicialment que són el tipus d'usuari, el tipus de representació del Sistema d'Informació i el tipus de dades i, per tant, la pròpia naturalesa de les

dades.

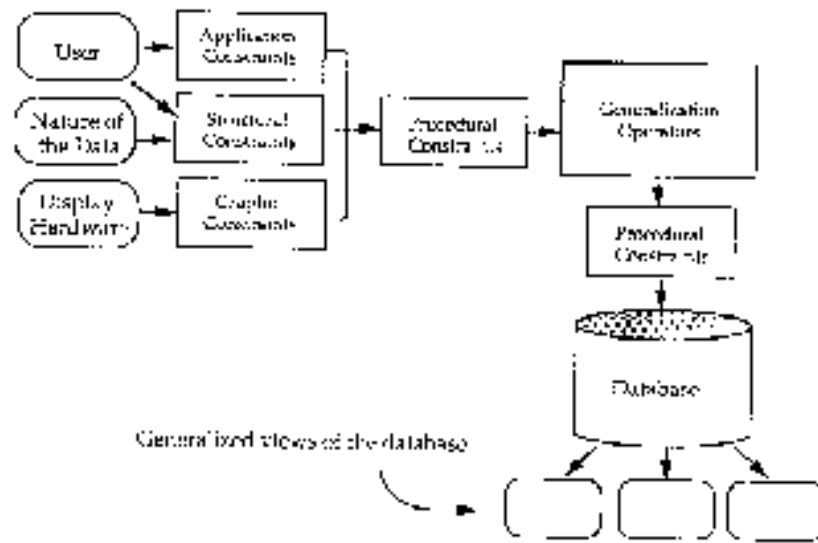


Fig. 7.3 Refining the constraint-based framework

Marc conceptual de treball proposat per Beard equivalent al de Brassel i Weibel on les limitacions són sustutuides per regles més flexibles del tipus If: Then (Beard 1991:123)

Un punt crític pel bon funcionament d'un sistema d'informació espacial és la necessitat de permetre una relació amb l'usuari amb un llenguatge d'alt-nivell que pugui matisar els continguts de les preguntes que es planteja l'usuari a un nivell superior del que existeix en l'actualitat. Nyerges evidencia com, degut al fet de que la representació digital de la realitat correspongui a la base de dades geogràfica,

“We need to extent our treatment of geographical database to include knowledge of geographical meaning. Data management capabilities must be enhanced with representation capabilities. [...] Knowledge of geographical meaning as part of the abstraction process is seldom captured since database development of primitive points, lines and areas with attendant attributes has been a tremendous job in itself. One of the reasons why experienced cartographers have not fully described how generalization is done is because they have never had a means to systematically document the knowledge they use to perform a generalization task. The fundamental of that knowledge concern meaning. Meaning involves content as well as a definitional

ontology of geographical structure and process. Using a high-level somewhat formalized language that is probably the best way to proceed examining the nature of the map generalization problem." (Nyerges 1991:59)

Hi ha diverses possibilitats en els sistemes d'informació geogràfica que ens poden permetre no haver d'interaccionar directament amb les funcions de baix-nivell sinó de fer-ho a través d'una *shell* que transfereixi la intencionalitat de la tasca a desenvolupar i presenti un esquema de fluxos, sense que l'usuari hagi de conèixer quin tipus de model de dades hi ha implementat, sinó que la interfície permeti un diàleg basat en els objectius i tasques i no en les funcions primitives associades als diferents models de les dades. Un dels esforços en l'establiment d'un model que no estigui directament lligat a l'estructura de les dades és el presentat per Tomlin que desenvolupa un llenguatge d'alt-nivell (Tomlin 1990) que, encara que s'inspira principalment en el model geomètric de dades *raster*, es pot considerar independent de la geometria i utilitza una seqüència lògica de funcions de manipulació per resoldre problemes espacials.

És, doncs, necessari poder plantejar les tasques a través d'altres comandaments que els actuals primitius dels sistemes d'informació geogràfica, ja que molts errors provenen del desconeixement del que s'està fent i de quins són els efectes que es poden originar amb cada funció o comandaments. L'usuari té un model mental de la manera de com ha de realitzar una seguit de tasques que l'han de dur a poder prendre una decisió, i la conducta humana es basa en l'establiment d'uns objectius que ens indiquen quan la tasca s'acaba. Aquests objectius els podem descriure com accions o estats que s'han d'aconseguir a través de l'aplicació d'una sèrie d'operadors (Albrecht 1994:212).

Però actualment en els SIG es deixa en mans de l'usuari el coneixement de totes les possibles funcions del sistema que s'està utilitzant. El fet que les funcions siguin específiques dels models geomètrics de les dades (*raster* o vectorial), fa que tot el posterior tractament de la informació depengui del model conceptual del món real que hagi establert l'usuari. La poca correspondència en el programari entre les tasques i les operacions que ofereix el model geomètric per manipular funcions contínues o discretes i camps o objectes, fa que s'introdueixin molts errors de tipus conceptual a

traves de la mala utilització dels comandaments de baix-nivell, que és l'eina de relació amb el sistema que té l'usuari.

És important intentar facilitar la transparència entre l'interacció en termes del mateix problema i no de les operacions dels GIS. Una possible solució seria poder comptar amb un GIS virtual, ja que permetria l'enquadrament del tipus de tractament necessari a partir d'una declaració de les dades introduïdes i d'una caracterització de les dades segons les necessitats del tema, format i el grau de correcció (*accuracy*) en funció del domini on es defineix el problema. I a continuació aquest GIS virtual podria definir una seqüència de procediments correctes sintàcticament, necessaris per la conversió de les dades en informació i poder seguidament identificar quin es el millor model de dades segons l'esquema de transformació d'Albrecht (1994:214), a partir de Kemp (1992).

Un VGIS no és altra cosa que un sistema basat en el model cartogràfic "invers" de Tomlin, ja que intenta, a partir de les funcions de baix nivell, trobar el millor model de dades per un camp d'utilització. Així ho assenyala Albrecht que assenyala que el model cartogràfic és un llenguatge d'alt nivell independent de l'estructura que el model de les dades tenen.

"[Unfortunately] there is little correspondance between the user's tasks and the functions offered by GIS. [...] Specifically what is missing is a high level language as an interface that would allow the user to declare the problem, allow some interpreter to evaluate it, and generate a solution composed of the Gis command language. In short, the interaction with a GIS would be in terms of the problem domain, not primitive GIS operations"

"[...]VGIS is seen as a first step towards an automatic use of the immense knowledge that is inherent to a GIS. It models generic objects (visualization). The user describes the target information within a shell that is placed on top of whatever GIS. [...] No matter what task the users want to accomplish, they have to deal with whatever data ther is.[...] eventually, necessary conversions will be performed an utomatically documented within the framework of task-driven user interaction." (Albrecht 1994:215)

Reconeixement de l'estructura: *informació sobre les dades*

Ja hem vist en el capítol 4 que dins del procés de generalització el preprocessament és on es defineix l'esquema d'integració de diferents conjunts de dades i com es realitza a nivell conceptual d'integració en base a un dels elements

funcionals com proposa Muller, limitatius com proposa Beard, o de control com ho fa Robinson, que és l'escala.

Identificar les propietats i relacions entre objectes que necessitem per la generalització, ho fem en base a les característiques de tipus geomètric, topològic o relacions espacials i propietats semàntiques.

1) El control de la geocodificació en la recollida de les dades (De Man 1984). La tria del sistema adequat de geocodificació ha de tenir en compte la variabilitat espacial dels fenòmens, les possibilitats de recollida de dades i el context de decisió. L'estructura de dades espacial, o sigui, la manera de georeferenciar les dades i emmagatzemar-les digitalment, és diferent segons es tracti de dades satèl·lit de tipus raster (que difereixen de totes les altres per l'alt grau de correcció, una també alta resolució espacial i temporal i un baix nivell d'abstracció humana) o les provinents de mapes amb punt i línia (que presenten una sèrie de fenòmens seleccionats de manera altament abstracta i generalitzada).

El sistema de referenciació geogràfica o sistema de geocodificació consisteix en l'assignació d'un fenomen a una referència espacial, i el mètode que triem de codificació dependrà de les necessitats en funció del tipus d'anàlisi que desitgem fer. El sistema de codificació del pas de la realitat a la definició d'un tipus d'estructura que en la cartografia tradicional es fa de manera automàtica passant de dada a codi segons el "rapid prototyping" no és un procés gaire ben estructurat.

El tamany òptim de les unitats bàsiques espacials d'un conjunt de dades depèn de la precisió del fenomen que considerem. En general, hem de poder imaginar que el nivell òptim de detall és un acord entre les necessitats d'anàlisi dels diferents conjunts de dades, i de tamany més petit que qualsevol de les unitats dels diferents conjunts de dades.

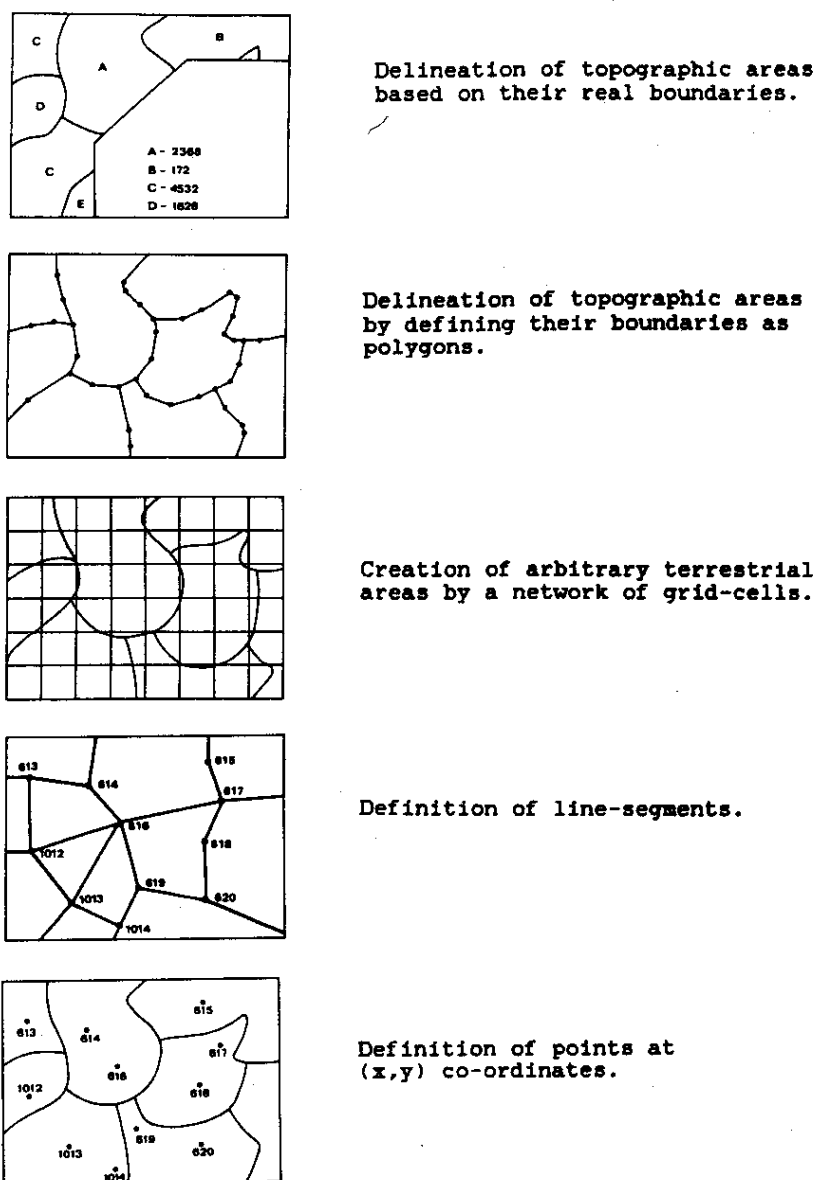


Fig. 4: Different methods of geographic referencing (after Tomlinson, 1972).

(Font: Man 1984:33)

El nivell de detall d'una representació cartogràfica ve definida en funció de la finalitat. Abans de començar a muntar un sistema d'informació geogràfica que permeti la integració de dades de tipus heterogeni hem de definir el tipus d'utilització que en farem de la mateixa manera que, en la cartografia clàssica hem de tenir clar quina és la finalitat i el tipus d'utilització volguda. Això comporta decidir sobre l'"accuracy" necessària, la precisió i la resolució que necessitem en el nostre anàlisi posterior.

2) La integració de diferents dades espacials la fem en base al tipus d'objecte espacial i no al de la seva representació cartogràfica. Un mètode apropiat de referenciació geogràfica es pot triar una vegada hem definit el tipus de tractament de les dades espacials, el nivell desitjat de detall i la necessitat d'integrar diferents conjunts d'informació espacial.

La integració de diferents conjunts de dades espacials la podem fer aplicant un mateix sistema de codificació en el moment inicial de recollida de les dades, o bé trobant mètodes que permetin la transformació d'un conjunt de dades en l'esquema de l'altre conjunt de dades, que podem diferenciar com a "bàsic" o "derivat".

El concepte de variabilitat (*high degree on variation = low decay rate*) suposa una component dinàmica que és important al comparar diferents conjunts de dades per la seva integració, ja que tots els conjunts de dades es caracteritzen per tenir una variabilitat que va des de valors alts a baixos. Els conjunts de dades procedents de fenòmens naturals, recollides a partir de fotografia aèria i imatge satèl·lit, posseeixen en principi una variabilitat menor que els de tipus socioeconòmic. Podem utilitzar aquesta informació amb altres provinents de mapes analògics sempre que estiguin a una mateixa escala per augmentar-ne el nivell de detall i l'"accuracy" en el moment de classificar.

El grau de correcció és citat per Croswell com un concepte ambigu per la dificultat en establir la relació existent entre l'escala en la compilació original de les dades, la correcció semàntica i l'adequació a la seva correcció mètrica.

"Map accuracy is dependent upon the scale of map compilation. There is a direct relationship between map scale and positional accuracy that can potentially be depicted. (Croswell 1988:450)

Podem considerar com a norma general que els diferents tipus de dades provinents dels recursos naturals que tenen un alt grau de variació en l'espai i un baix nivell de *decay rate* són apropiats per ser considerats dades "bàsiques" en un sistema integrat d'informació.

És important considerar que la qualitat d'un conjunt de dades es pot caracteritzar pel seu nivell de generalització en termes d'escala o *decay rate* o variabilitat, i és aquesta característica la que ens pot donar informació sobre quines són les limitacions de les dades.

Definir les restriccions o limitacions: el nivell òptim de detall espacial i el mètode de referenciació les triem en funció de les necessitats d'anàlisi

Els processos de transformació i modelització del procés de generalització es poden considerar un procés dirigit a aconseguir maximitzar el contingut de la informació (information-oriented) a la vegada que respectant una sèrie de limitacions.

Una limitació restringeix les alternatives que estem considerant, i es deriven dels controls citats per Robinsonn Robinson et al. 1984 i mes recentment per Brassel i Weibel Brassel, Weibel 1988 : el propòsit del mapa, l'escala, la naturalesa i qualitat de les dades i les limitacions gràfiques⁷

Els controls són els encarregats durant el procés cartogràfic de definir un conjunt de limitacions, i que en la "vista" de la base de dades, són les limitacions les que s'encarreguen a la seva vegada de satisfer un conjunt d'accions, que han de o bé preservar les condicions inicials, quan es compleixen les limitacions o bé han de poder corregir les condicions si les limitacions es veuen alterades. El resultat final és una "vista" generalitzada de la base de dades on totes les limitacions es troben especificades.

Beard identifica les limitacions que tenen sentit en la generalització així com la seva interacció, i les organitza segons els tipus de saber d'Armstrong (1991) en: limitacions de tipus gràfic, estructural, d'aplicació i processuals. Assenyala que en l'automatització del procés de generalització l'establiment de regles es difícil, i proposa que aquestes es canviïn per limitacions (*constraints*).

"Sensitivity to spatial and attribute variation, to map purpose, and map scale suggests the need for flexible or modifiable rules. Standard rule-based systems use the syntax IF (predicate) : THEN (consequence) with the predicate being a condition or

⁷ La finalitat de la generalització per el que respecta a la modelització ha de comptar sempre amb la representació gràfica. El model de les dades o DLM esta organitzat i es generalitza de manera no innòcua per el que fa referència al procés posterior de conversió al model cartogràfic.

Per una altre banda no ha d'oblidar que que la qualitat del producte cartogràfic esta relacionat a l'adequació (fitness) respecte a la seva utilitat. Per tant totes les operacions de manipulació de les dades hauran d'esser controlades.

combination of proposition about the database and the consequent, a collection of actions. It is therefore not an optimal foundation for automating generalization.[...] an alternate approach is proposed to accommodate flexibility [...] constraints substitute for rules. A constraint is a condition similar to the predicate in a production rule. The distinction is that a constraint is not bound to a particular action. The overall rule is that all constraints" (Beard 1991:123)

De la mateixa manera que en la cartografia, on la confecció d'un mapa s'ha fet sempre a partir de dades a escales més grans, en el model de dades necessitem derivar a través de remostratge de les dades originals les dades que derivem en un model de dades derivat, en funció del tipus d'anàlisi que precisem. Tobler assenyala com:

"Spatial resolution in the Earth sciences has long been recognized as being closely related to the sampling interval. Thus the meteorologist knows that a geographically dense net of observing stations is required to track the fine details of a thunderstorm. These storms are generally from one to five km across. [...] Since observations are never perfect, the better rule of thumb is to use a sampling interval one fifth the size of the feature to be detected. This has an important implication for geographical data 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 to a collection of geographic data to be suitable for all kinds of problems. Furthermore, one must recognize the existence of invisible phenomena and must also recognize the aliasing effects which result from the spatial size of the sampling interval." (Tobler 1988:134)

Com diu Tobler, no és possible disposar de dades que siguin utilitzables per qualsevol aplicació. És necessari poder-les derivar, tal com es fa en la teledetecció. Així doncs les possibilitats de les que disposem consisteixen bé en disposar de sistemes jeràrquics que permetin la integració dels models a través de les seves definicions en un diccionari, a través dels atributs.

Definir les possibilitats de canvi: *un bon disseny de la base de dades que permeti una integració de les "vistes"*

La manca d'un organisme central estatal que coordini la recollida de dades ha

fet que en molts països aquestes s'hagin recollit sense contemplar les necessitats de la seva integració. Nyerges cita com a principal obstacle la manca d'una metodologia que permeti la integració de la informació i la possibilitat de compartir les dades. O sigui que no consisteix en considerar les dades com un tot sinó en fer que diferents parts cooperin. La metodologia per integrar informació fa referència a la integració de: dades, programaris, maquinari i/o persones.

La integració de les dades es pot fer a nivell de les dades de manera aïllada o de la base de dades de manera més global. Dos tipus de dades són importants en la integració de les dades: les descripcions que es troben en un esquema i les descripcions o definicions que es troben en un diccionari. Com que les descripcions estructurals de l'esquema descriuen la manera com les dades es representen a la base de dades, i és diferent per cada objecte espacial, les descripcions han de formar part de l'esquema. En canvi, les descripcions de les dades del diccionari proporcionen una interpretació semàntica de les entitats de l'esquema, a més de donar definicions de les descripcions espacials, d'atribut i físiques. Són aquests esquemes i diccionaris de les dades les que s'anomenen metadades o dades sobre les dades, i la integració de la informació té com a finalitat la integració d'aquestes. La metodologia es desenvolupa en quatre fases: pla estratègic d'integració, anàlisi de la integració, disseny de la integració i implementació de la integració; cada una de les fases es desenvolupa a tres nivells. Aquests tres nivells són: comprensió conceptual; les tècniques que s'utilitzen per expressar els conceptes en termes o *data constructs*; i el programari/maquinari per la implementació d'aquestes construccions. I el resultat és un conjunt d'objectius de la integració, un conjunt d'especificacions de "quina" informació i contingut existeix o ha d'existir; un disseny de "com" s'ha d'emmagatzemar el contingut d'aquesta informació; la manera d'aplicar el pla d'integració.

Com assenyala Nyerges, la dificultat major consisteix en la dificultat de comparar si dues bases de dades tenen continguts compatibles ja que això suposa, a la vegada, tant compatibilitat estructural com locacional.

"The goal is to generate a specification for a single, compatible schema that can be used as the schema for a land use information system"

Però a continuació assenyala que el que és un avantatge en l'esquema d'integració és la seva debilitat, ja que si oblidem una equivalència trenquem la

cadena.

"The long term goal [...] is to explore the influence of semantics on spatial structures and the requirements for spatial analytic models integrated into a GIS."
(Nyerges 1989b:30)

5.2.2.2. Formulació dels objectius: *integració de les dades durant el procés de compilació digital de la informació*

La identificació dels objectius permet facilitar l'estructuració de les regles de presa de decisió que es poden resumir en: el reconeixement de l'estructura i la identificació dels objectes cartogràfics, més les relacions espacials; la identificació de la dificultat en el tractament de l'escala a través de la seva substracció com escala conceptual i deixant només l'escala mètrica del procés, durant el procés d'integració per identificar quin tipus de limitació de tipus geogràfic, tècnic o conceptual correspon a la dificultat de l'usuari per determinar l'escala adequada d'anàlisi

Si considerem la distinció entre dos tipus d'objectius el que es necessari i el que seria convenient (*soft*), podem considerar com un objectiu *hard* la integració en base als formats de les dades i la seva conversió a partir de la integració dels nivell lògic o esquema conceptual, i a nivell conceptual baix dels nivells d'abstracció; i l'altre objectiu de tipus *soft* que seria la consecució de la integració a partir de un nivell més abstracte dels nivells d'abstracció, a partir de la integració dels models conceptuais. Podríem dir que el primer correspondria a la integració a partir de les resolucions de les dades i el segon es faria a partir de la integració dels models de les bases de dades, tal com es presentat per Nyerges i que queda recollit en el capítol 4.

Podem dir que actualment coneixem gran part de les resolucions, principalment però no les escales de les dades ¿hi ha una relació entre resolució i nivell de detall i escala (*level of geographic detail*), ([Goodchild, 1997 #545]:2) de representació o d'anàlisi?

En realitat es pot representar la incomprensió de quina és la vinculació possible entre l'escala conceptual i la resolució ja que ha vingut representada per una tendència, en el cas de la resolució, més tendent a creure en la transformació automàtica entre escales que no en la impossibilitat de relacionar certes resolucions a unes escales i que entre elles no permeten tenen una continuïtat.

La generalització, diferent de la vista del model o de la vista cartogràfica (*cartographic view*), permet considerar des d'una perspectiva diferent el procés de compilació, la manipulació prèvia de la base de dades en la generalització del model. Si recordem que les raons que ens interessin de la generalització en el model de les dades hauria de facilitar

- l'accés a les dades en els SIG, on l'usuari pugui manipular les dades sense haver de conèixer la seva representació interna

- permetre que les consultes analítiques dirigeixin el procés, com seria el permetre conèixer quina és la mitjana espacial o les noves classes que han d'aparèixer a un nivell determinat de variança

En aquest punt Brassel i Weibel es pregunten:

"How much and what kind of generalization support is required for the accomplishment of routine tasks in cartographic generalization?" ([Muller, 1995 #127]:6):

Les tres principals raons per automatitzar la generalització en un sistema d'informació geogràfica són les següents:

- el crear un nivell de detall apropiat per la escala d'anàlisi
- el permetre l'anàlisi de les dades a diferents nivells de resolució
- i el minimitzar l'emmagatzematge i les operacions de entrada/sortida

Les dues primeres correspondrien als principals problemes que trobem en la integració de dades, i a les que ens referim en aquest treball.

Objectius en conflicte:

- ***hard: integració a baix nivell o resolució;***
- ***soft: integració a alt nivell o escala conceptual***

La tensió que presenta la generalització en un entorn digital ve representada per la diferent concepció d'aquesta en l'entorn digital més centrada en la forma geomètrica o en l'objecte cartogràfic. Aquesta tendència ha vingut representada pel tractament de la forma geomètrica amb l'aplicació d'algoritmes, operadors o funcions de tractament purament gràfic. L'intent d'aplicació de fractals s'ha generat en aquesta

creença de tractament geomètric i ha estat durant uns anys una possibilitat seguida amb més fe que altra cosa ja que ha intentat suplir el nivell superior del “sentit” de les formes cartogràfiques. La recuperació de la generalització com a procés més holista i global recull el que van exposar escoles anteriors, tant a nivell de tractament quantitatiu de les dades com de tractament gràfic dels anys 60’ i 70’.

Brassel fa un bon resum de quina ha estat la generalització conceptual en els inicis en l’entorn digital. En els anys 60’ i 70’ l’escola quantitativa introdueix diversos mètodes i programes per la classificació de les variable temàtiques. Exemple d’això és l’intent de definició (*verfahren*) i agregació (*zuberechnung*) de les fronteres de classes en els mapes de coropletes i isopletes, o en la classificació dels mapes nominals. És bastant evident que l’agregació d’objectes discrets nominals en classes superiors és sempre factible quan es coneix l’estructura, com acostuma a ser possible en el cas de les ciències naturals.. A partir dels anys seixanta s’utilitzen també mètodes multivariants per decidir la pertinença de determinats objectes a classes nominals que tenen més d’una característica quantitativa. Exemple d’això és la tipificació de les unitats administratives per les necessitats d’anàlisi, (anàlisi factorial, anàlisi de *clústers*, anàlisi discriminant) de varies variables socioeconòmiques. Més exemples es trobarien en la construcció de base de dades jeràrquiques on els objectes es codifiquen segons que siguin punt línia o àrea en el seu emmagatzemament, que es indicatiu de l’objecte.

Es tracta, doncs, de poder comptar amb l’ajut en l’entorn digital també en les tasques de selecció, que en l’actualitat corresponen a una decisió de l’usuari que és el que codifica la informació.

Morrisson assenyalava que en el mapa analògic tradicional el grau de precisió mètrica i l’escala a la que es visualitzava la informació eren elements que estaven en conflicte (Morrison 1994:7).

Crec que hi ha d’entrada un conflicte entre l’usuari actual que ha de poder tenir coneixements del seu camp, amb uns coneixements amb una tradició de tractament espacial de la informació com és la geografia, i amb uns coneixements de comunicació del cartògraf. Això és difícil de substituir per un *software* d’un programa per molt “saber” que “utilitzi”.

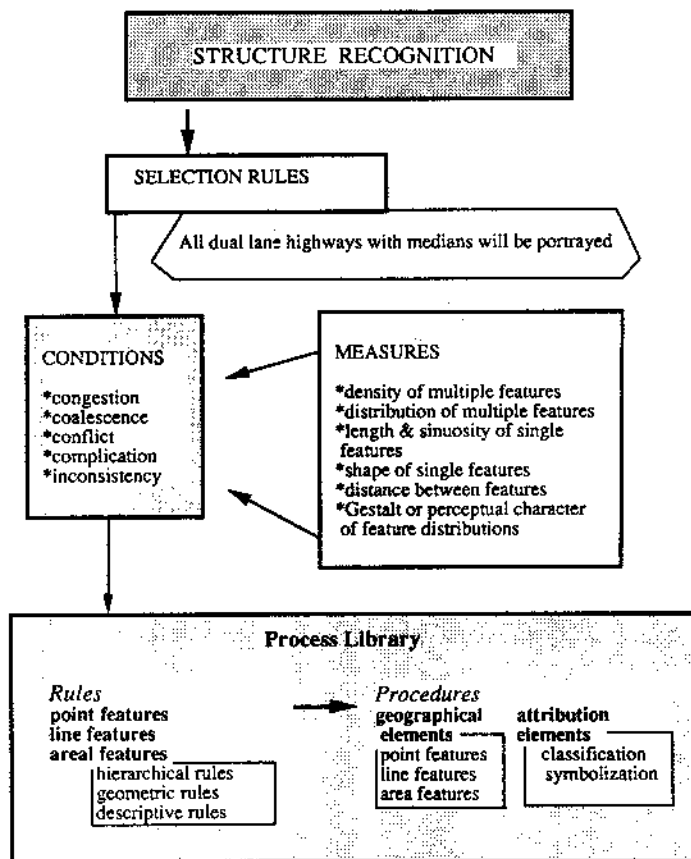
An objective is thus a perspective that serves to guide the structuring of decision rules. I distingeix el terme objectiu com a sinònim de funció de l'objectiu en programació matemàtica i trïa de la funció com s'utilitza Eastman et alt. (Eastman et al. 1995:4).

Unes regles de decisió s'estructuren en base a un objectiu i els diferents criteris. La naturalesa de l'objectiu i com es veu per part del *decision maker* té una forta influència sobre la regla de decisió a seguir.

***Criteris que representen el més específic en la jerarquia dels objectius:
variabilitat espacial***

Brassel assenyala que els processos de la generalització estadística són guiats per criteris quantitatius, i no són utilitzables posteriorment en la generalització cartogràfica (Brassel, Weibel 1988:229).

L'escala, que té sentit en el mapa en paper perd sentit en una pantalla on podem fer *zooms* endavant i endarrere. El que sí que ha quedat és l'escala d'anàlisi dels objectes geogràfics. Partir de bases de dades on l'escala és desconeguda fa que es perdi, de cara a l'usuari, una tradició analítica basada en l'escala a la que succeeixen certs fenòmens. La resolució es converteix en una substitució de l'escala en els nous entorns ja que representa unes característiques de precisió en la recollida. Però el que és evident és que el procés es realitza a un nivell conceptual més abstracte.



(Font: Brassel, Weibel 1988)

En el quadre corresponent al procés de generalització en un entorn digital es proposa un procés basat en el de Brassel i Weibel. Donades les informacions de partida un primer pas consistiria en el reconeixement de l'estructura (a) que consisteix en la identificació dels objectes o agregats dels objectes, quines són les necessitats espacials i quina és la seva importància. Després de l'anàlisi de la informació inicial el procés de generalització (b) s'ha de definir, haurem de definir el procés de manipulació i els paràmetres de les dades finals. A continuació els processos els podem tenir en una biblioteca de processos; per tant veiem que els processos depenen del reconeixement de les dades inicials, que han de permetre la confecció del mapa final.

Ens pot ajudar comparar el procés de generalització cartogràfica per veure com l'estadi inicial de selecció dels fenòmens i el pas al model de les dades és decisiu en el procés posterior, tant en la generalització cartogràfica com en la manual.

5.2.2.3 Modelització de la situació desitjable

Fer el procés de generalització i d'integració en base a la comprensió és la proposta de Brassel i Weibel, i també de Nyerges, tal com ja s'ha assenyalat anteriorment en els capítols 2 i 4. Això suposa poder comptar amb els dos primers estadis de reconeixement de l'estructura i reconeixement del procés, d'una manera automàtica o amb coneixement (a alt nivell), per derivar a partir d'una escala gran una de major que presenti un major grau d'abstracció: "some guiding knowledge at a higher conceptual level will be necessary to control the application of individual algorithms to different situations" (Weibel 1991:174).

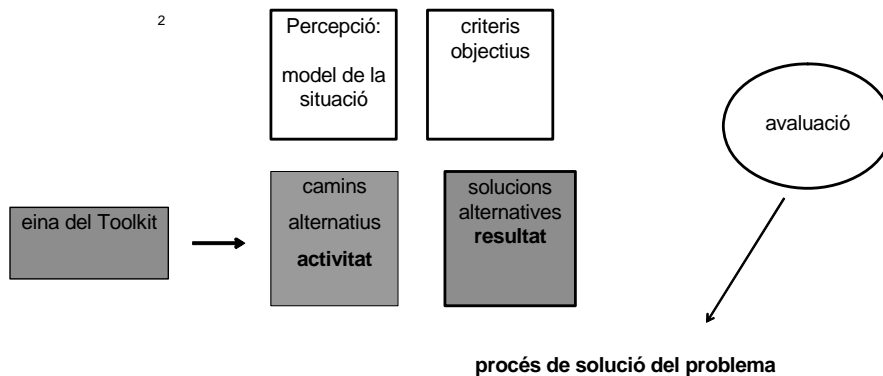
Com diuen Brassel i Weibel, seria desitjable poder comptar amb fer l'anàlisi de les dades a múltiples nivells de resolució: "an ideal generalització procedure within a gis would select the best available generalització process based on purpose, scale, data types and data quality requirements" Joao 1991:493. El mateix és assenyalat per Mark basant-se en el tipus de fenomen en el món real (Mark 1991:103).

Ja l'any 1989, Feuchtwanger resumia com s'havia de realitzar el disseny d'una base de dades mitjançant quatre preguntes (Feuchtwanger 1989:600):

- 1) quins són els principals tipus de models de base de dades?
- 2) qui utilitza el model de la base de dades?
- 3) quines són les principals "vistes" d'una base de dades geogràfica?
- 4) quins són els principals dominis d'una base de dades geogràfica?

5.2.2.4 Camins alternatius per crear solucions

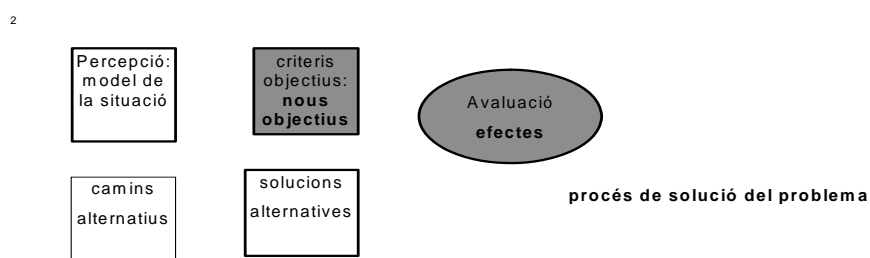
D'entre les solucions alternatives, cal destacar la recuperació d'antigues tècniques tant analítiques com de representació que, amb la possibilitat de manipulació digital, són factibles. Un exemple n'és la conversió de les dades en superfícies estadístiques (ja començat per Imhof) per la seva integració a través de la interpolació, o coeficient de covariació a les dades digitals de la base de dades.



5.2.2.5 Percepció del model de la situació resultant

És necessari poder comptar amb una percepció del model de la situació resultant per tal de prioritzar una visió orientada al model i a l'anàlisi

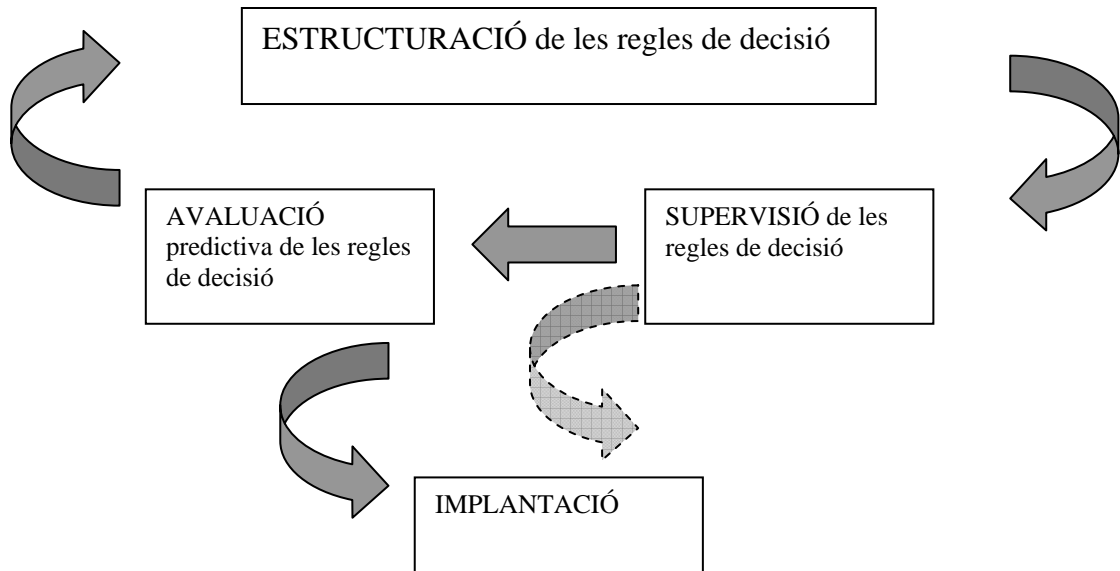
"The traditional view of generalization in support of surveying and mapping has neglected model-oriented generalization [...] busy in implementing algorithms to perform the analogue of cartographic generalization have forgotten the intimate relation between generalization at the modelling level and generalization at the surface (cartographical representation)



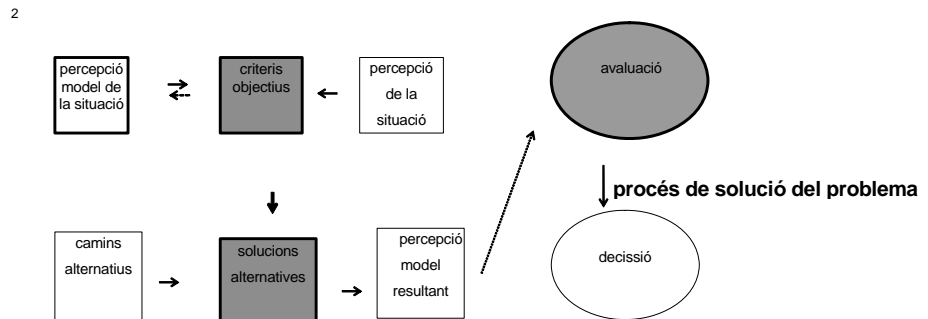
5.2.2.6 Avaluació

Brassel i Weibel fan explícit que, en relació amb la generalització estadística hi ha el problema dels errors provinents de la base de dades cartogràfica i la propagació d'aquest durant el procés de manipulació, com ara en la sobreposició de capes

temàtiques (*overlay*), cosa a la que en l'actualitat es presta molta atenció (Brassel, Weibel 1988:236).

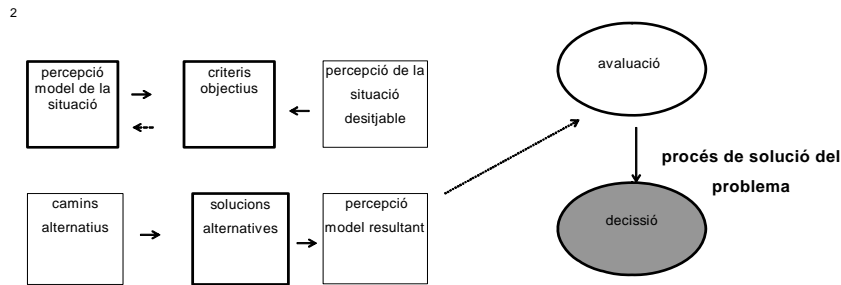


O bé representat d'altre manera:



5.2.2.8 decisió

És en el preprocessament de les dades en un entorn digital on es decideix la precisió amb la que es treballarà, i en general la qualitat del procés que es desitja i el tipus de objecte amb els que volem realitzar l'anàlisi.



- Qualitat de les dades, qualitat del procés i de la solució

La generalització afecta algun dels quatre components: *accuracy* locacional, *accuracy* dels atributs, *consistency*, *completeness*. Hem de poder oblidar el problema de la representació i variar el nivell de precisió a la que els objectes espacials han sigut enregistrats en la base de dades. "Asking the scale of a database is in fact making a query about the lineage of the database". La generalització tant cartogràfica com estadística poden afectar la *accuracy* de les bases de dades espacials.

En la generalització estadística o filtratge que té com a finalitat la modelització espacial de la informació d'atributs lligada a la localització. Son constreyniments necessaris durant el procés de generalització estadística: la preservació de la mitjana, la variança i la forma de la distribució espacial. Així com la classificació errònia pot amagar dades significatives.

En canvi, la generalització cartogràfica pot afectar en gran manera tant la localització dels fenòmens com la seva forma. S'han aplicat criteris per estimar l'*accuracy* de la generalització cartogràfica. Els objectes que son self-similar no haurien de quedar afectats per la generalització. Així mateix els productes provinents de una digitalització i es tendeix a tractar-los com mes acurats del que en realitat son. Ho son menys que els derivats de un DLM.

En el cas de bases de dades, la resolució espacial (habilitat de un sistema de registrar i poder distingir objectes que estan a prop entre si) esta determinada per la qualitat de la mostra que va ésser realitzada per el recollidor. Una mostra *fina* pot preveure la millor resolució que una mes *cruda* , però les dades s'han de resampling a nivells mes grollers de resolució. També ho fa la generalització incrementant el

tamany mínim de les àrees. A mida que es reueix la resolució alguns fenòmens que eren visibles desapareixen mentre que altres apareixen . “A coarser resolution or a higher level of generalització may provide more explanation over the variance of a spatial variable than finer resolution levels (Tobler and Moellering 1972) (gis voll 1991:459 Muller)

- Qualitat de les dades: *metadades*

El processament de les dades, que podem controlar a través de un llinatge permet controlar les dades, pel que fa a la qualitat, informació sobre aquestes i quines possibilitats ofereixen per aconseguir els propòsits que ens hem proposat.

Dins d'aquest apartat tindria una gran importància tot el que fa referència a les dades ja que es un camp que genera i generarà molts diners, l'escut del preu es la seva qualitat, i per tant es tendirà a vendre dades per sobre de les necessitats de accuracy i precisió necessàries per l'usuari. L'escut de la qualitat de les dades es basa en la pressuposició de que unes dades bones duen a bons resultats, i això no es ben be cert. El que si es cert es el fet de que dades “dolentes” porten a resultats erronis.

Com a conseqüència de la manca en les bases de dades actualment de la informació/coneixement per a poder generalitzar, hi ha un esforç en la recerca actualment, en el preprocessament, pas que consisteix en derivar o afegir informació addicional i que és considerada freqüentment com a “metadata” (Lagrange & Ruas, Gisdata, 1994:1100)

La recollida de dades es un dels costos mes grans el poder aprofitar les dades entre organismes fa necessari uns estàndards de qualitat i de transferència de cara ala seva integració en altres bases de dades. Podem dir que les metadata o “dades sobre les dades” o “informació que donen sentit a les dades”. El que es cert es que s'els dona diferents significats des de format fins els processos. Els que ho entenen en l'àmbit dels metamodels, o sigui, des de les estructures conceptuals subjacents, veuen les metadata com les dades més bàsiques.

Des de un punt de vista funcional per (D. Medyckyj-Scott et alt 1996:66 IJGIS) podem considerar com un recurs :

- en el data discovery
- el transferiment de dades

- en el ús de les dades i la seva interpretació

I podríem resumir-ho com informació que dona meaning a les dades, tot i sabent que el meaning depèn del context d'ús.

La gestió de les dades i us de metadata per projectes individuals o petits projectes han sigut discutides per altres entre el que hi ha Lanter que té un programa de “rentat de les dades” anomenat Geolineus.

- Qualitat del procés: *metallinatge*

Un dels grans inconvenients en el control del canvis que es van fent a les dades o a les seves representacions es la falta de un procediment que permeti fer un seguiment de quins han sigut els passos. Joao (Joao 1991:493) assenyala el fet de els GIS no permeten normalment fer un seguiment i per tant tampoc permeten doncs tornar endarrera en el procés.

El procés de generalització no es inoqu pel que fa a l'acuramenta dels productes cartogràfics, ja que introdueix efectes imprevisibles en la mètrica, topologia i semàntica del producte cartogràfic. Els canvis que s'introdueixen durant la generalització com poden ésser el lateral shift o la distorsió angular afecten l'anàlisi en el GIS durant el solapament cartogràfic Joao et al. 1993.

“Corresponding criteria to judge the performance of map generalization methods were exclusively geometric. Those methods were limited to isolated objects taken out of context and were mainly applied to large scale topographic maps. They emphasize the HOW rather than the WHEN, WHAT, and WHY. (Shea and mcmaster1989; Richardson 1988).In so doing, automated generalization was breaking away from the long tradition of manual generalization where geographical analysis and evaluation of map content were constant guiding inputs to the process. Recent research efforts have demonstrated the need for an expanded, more integrated approach to the problem of automated generalization (Buttenfield and delotto 1989). “ (Richardson, Muller 1991:136 MG)

Brassel and Weibel descriuen la generalització cartogràfica relacionant-la a la reducció d'escala en el mapa mentre que s'intenta preservar lo representatiu del mapa de manera holística, de manera no estadística. Degut a que el procés es realitza sota control estadístic, assumeixen que les diferències existents entre la base de dades geogràfica original i la derivada son mínimes i tenen poca influència en la base de

dades cartogràfica. Nyerges creu que les tècniques estadístiques per realitzar les operacions no son suficients per controlar la pèrdua de “meaning”*.

Reichenbacher cita el concepte Armstrong, Bennet (1990 Gis/Lis) de metaplanner que seria el sistema que ens permet plantejar un metode de resolució: ”ine metode zur verfügung stellen, die einen Plan oder mehrerepläne entwirft” ([Reichenbacher, 1996 #141]). Aixó és molt convenient ja que en problemes mal estructurats com és de la generalització l’usuari hauria de poder triar entre una serie d’alternatives quina creu que es la millor i en base a aixó presenta Armstrong el concepte de Spatial Decission Support System.

- Qualitat de la solució

La possibilitat de poder disposar de eines que ens permetin avaluar la qualitat dels resultats de la generalització no es encara un fet normal. No hi ha de moment encara criteris ni mètodes per fer-ho. Aquestes eines haurien de poder ajudar en la tria de un algoritme o la selecció de un d’exemples positius per eraining de xarxes neuronals. Si que existeixen tècniques d’avaluació parcials però no n’hi ha cap que pugui conside la interacció que es dona en la integració de diferents tasques. Hi ha una proposta de Mackaness (11 MG:226), que es troba en el inici de la línia de recerca, molt simple, on a través de la representació gràfica del diagrama d’una rosa permet comparar diferents tècniques , i defineix el concepte de topogeneralització que prové de topobiologia.

Els objectius en un nou entorn els sistemes d’informació geogràfica pel que fa referència a la generalització ens porta a reconsiderar els objectius i deslligar objectius i processos . Evidenciar que son els processos els que han d’haver canviat i no els objectius de la generalització , em porta a mostrar com és en el DLM on es plasmen els requisits que ens permeten aconseguir els objectius.