

Structuration of Phenotypes / Genotypes through Galois Lattices and Implications

Vincent Duquenne^{1,4}, Caroline Chabert¹, Améziane Cherfouh¹,
Jean-Maurice Delabar², Anne-Lise Doyen¹, and Douglas Pickering³

¹ FRE 2134, Institut de Transgénose, 3B rue de la Férollerie, F-45071 Orléans Cedex 02

² UMR 8602, Lab. de Biochimie Génétique, 156 rue de Vaugirard, F-75730 Paris Cedex 15

³ Department of Mathematics, Brandon University, Manitoba, Canada R7A 6A9

⁴ To whom correspondence should be addressed at FRE 2134

v.duquenne@wanadoo.fr

Abstract. The *Galois Lattice* of a *binary relation* formalizes it as a *concept system*, dually ordered in "*extension*" / "*intension*". All implications between conjunctions of properties holding in it are summarized by a (recursive) *canonical basis* -all basis having the same cardinality (see [MR #87k:08009]). We report here how these *tools* structure *phenotypes / genotypes* in behavior genetics. On a generic viewpoint, both situations comprise two binary data sets that are paired through either a column or a row matching, which raises specific questions. If the data are small, as compared with data bases in bioinformatics, this illustrates how these abstract tools can unfold and better interpretations.

1 Introduction

The outstanding developments of *data* and *knowledge bases* in bioinformatics raise the questions of "knowledge extracting" and putting data bases in "canonical forms" - in order to speed up access to information-, and of extracting classifications and rules providing some explanation on the biological topics under study.

On the other hand, it is now well established (see [1,12,15]) that the *Galois Lattice* of a *binary relation* formalizes any kind of *duality* -here between a set of *objects* (*subjects, patients...*) O and a set of *properties* (attributes...) P - and can be used as a general model for structuring it as a concept system that is dually ordered in "extension" / "intension": the lattice elements -sometimes called "concepts"- are simply the ordered pairs $(X,Y) \in O \times P$, where X (*extension*) is the *maximal* group of objects "having" Y , and reciprocally for Y (*intension*) which is the maximal subset of properties shared by X . The lattice elements are just ordered along their extensions.

This model was also used to enrich techniques and models that are implicitly based on *trees* (classifications, inheritance of properties...), and -through the equivalence between (finite) *lattices / closure operator*- could even be used as a general model for *formal languages* (see [5]) and for *evolving data bases* (see [2,3,24], and [14,16,17] for the algorithms). Lattices are more general and hopefully flexible than trees.

In a previous work (see [18, MR #87k:08009]) it has been shown that all the implications between conjunctions of properties holding in such binary data can be

summarized by a (recursive) *canonical basis* -and that all basis do have the same cardinality-, which can be expressed within the lattice through the existence of a minimal sub-structure -it's *meet-core* (see [10, MR #92g:06011])- made of *meet-irreducible* and *essential meet-reducible* elements on which the meet operation is restricted, out of which the whole lattice can be reconstructed -that generalizes G. Birkhoff's correspondence between *partial orders* and *distributive lattices*. Several papers have revisited these canonical form theorems (see for instance [15 §2]), and have either shown that the canonical implication basis could be helpful for contracting some classical results by D. Maier on data bases (see [6,22,23]), or have used them in statistics (see [9,13,20]) and in the analysis of symptoms in psychiatry (see [11,21]).

In this paper, we will take a more concrete approach and report on two current collaborations where these *lattice* and *implicational tools* have been used to structure the *phenotypes* and *genotypes* of subjects in behavior genetics (see [4,8]). Although the underlying data are in both cases quite small, as compared with the development of large data bases, it is worthwhile to illustrate how these abstract tools can unfold and better interpretations. On a technical viewpoint, both situations comprise two data sets that are paired through either a column or a row matching, respectively, which extends the usual situation with one binary table, and will raise new specific questions.

2 Implications for two groups of subjects: laterality questionnaire

For assessing handedness, several questionnaires are used on which multivariate analyses have been performed, showing some lack of agreement yet. Hence, the main goal of this report (see [8] for more interpretations) is to show how *lattice analysis* can help in understanding the associations among items, by comparing the results of left / right-handed writers for a questionnaire reporting which hand they use in life.

Right-handed writers. Basic data: a $S_{61} \times A_{11}$ 0/1-matrix where $(s,a)=1$ when the right-handed writer s uses *exclusively* his / her right hand for action a . Many subject's profiles are equal, and many right-handers use their right-hands for nearly all the actions. The profiles are ordered by (reverse) set inclusion -the closer they are to the bottom the more they are *consistent* right-handers- which is completed by set intersection that generates a *Galois lattice*. An element represents a *maximal* sub-table $S' \times A'$ filled with ones, for the subjects S' that share the actions A' , S' is the *extension* of A' , and dually A' is the *intension* of S' . The lattice is *minimally labeled*: an element's extension S' is restored by listing the subjects which are below it, and dually for A' .

Hence, the lattice is directly encoding the extension / intension *duality*, and gives an exact representation of which actions the subjects are sharing, and how this sharing is structured. The observed Galois lattice (simplified in Fig.1, see below) is small (190 el. / the potential $2^{11}=2048$ combinations), which reflects a strong structure of the right-handers / association of these actions. The actions are ordered by their extensions, which can be read as *implications* of which the *premises are single actions*. Hence, $T < r$ can be read $T \rightarrow r$: T :Throw implies r :racket. Similarly, B :Broom implies h :hammer, and C :Cards implies h :hammer and r :racket. Read from top to bottom, the lattice starts with $(s$:scissors, b :tooth-brush, m :matches, S :Shovel)

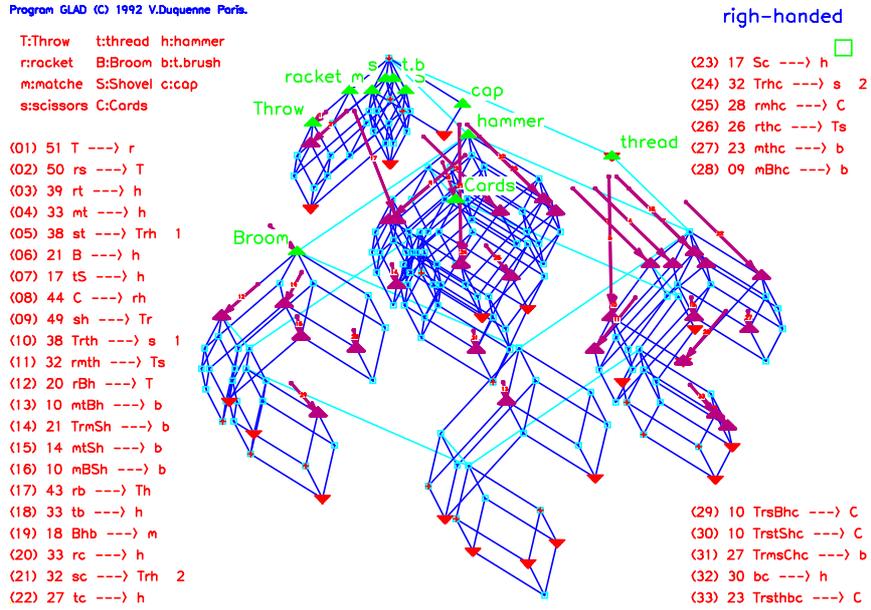


Fig. 1. Galois lattice and implications (right-handers)

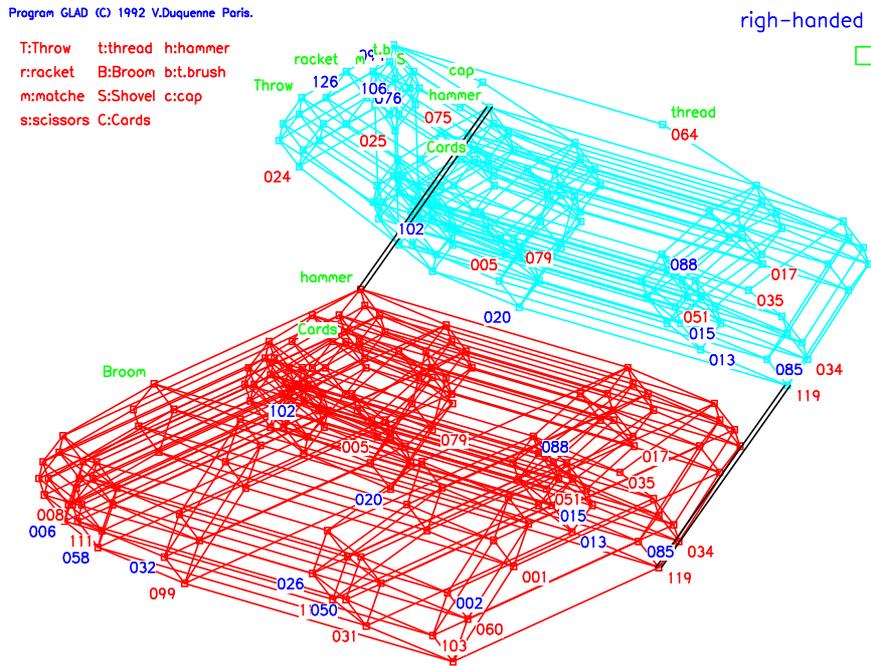


Fig. 2. Gluing decomposition in intervals (right-handers)

which are *independent* (generating a 2^*4 interval), then $r:racket$ and $T:Throw$ are independent with some of these, while $c:cap$ associates only with $m:matches$. All these actions associate more together when they are in conjunction with $h:hammer$. All the implications that hold can be summarized by a *canonical basis* (see Fig.1).

Those implications of which the *premises are conjunctions* reflect *discrepancy* to independence. They are weighted by the extension of their premise+conclusion's union. Most imply $h:hammer$ or $b:tooth-brush$. Fig.1 have been constructed by "blocking together" all the actions not below $h:hammer$, and then by introducing the remaining actions, to simplify the drawing by erasing some lines ("boxed" lattices in [15]). What is left by the action of implications is a kind of gluing of Boolean intervals which expresses (local) independence, which is extended by regular decreasing of the extensions' cardinalities. The lattice is decomposable by *un-gluing* (see [15], and Fig.2), which expresses *exchangeability properties* between actions that fall in the same intervals: they behave the same way for the two lattice operations which reflect the sharing of actions, and the intersections of the subject's groups.

Left-handed writers. Basic data: a $S_{23 \times A_{11}}$ 0/1-matrix, with the same set of actions. The profiles are nearly all distinct, except for one group. Here, $h:hammer$ and $C:Cards$ are equivalent, and both implied by $T:Throw$ (see Fig.3). Up to these implications, all the eleven actions but three are independent. On the other hand, $scissors$, $Shovel$, and $Broom$ don't associate much with all the other actions, due to the fact that the remaining implications comprise exactly one of these in their premises. Among the complex implications of the basis, most of them are satisfied by more than half the subjects, and several groups express equivalencies (#2-6 in Fig.3).

Out of the implications, $C \leftrightarrow h$ and $T \rightarrow Ch$ and the fact that $Broom$, $Shovel$ and $scissors$ don't associate with the other actions, the structure of the intervals reveals that the conjunctions of actions are mostly (locally) independent, that is also comforted by a regular decreasing of extensions. The left-handedness lattice is again decomposable in two intervals (see Fig.4), a decomposition also governed by $hammer$ ($Cards$) that concerns a majority of subjects (18/23). The lower interval is generated downwards by $Throw$, as compared with $Broom$ and $Cards$, for the group of right-handed subjects.

Comparison right / left-handed writers. The profiles of the left-handers are more diverse than for the right-handed subjects, which are somehow more stereotyped. In both populations $C:Cards$ implies $h:hammer$, while $T:Throw$ implies another action involving energy ($r:racket / h:hammer$), and it is interesting that both un-gluing are commended by $h:hammer$. The structure of left-handedness can be summarized by independence of eleven actions (up to the equivalence $C:Cards \leftrightarrow h:hammer$, which are both implied by $T:Throw$), together with non-association with the three remaining actions: $Broom$, $Shovel$ and $scissors$, which are thus sufficient to be quasi-consistent left-handers. The previous graphics give a clear picture on the hierarchies of actions for these populations taken apart, but it would be desirable to be in a position to characterize what is *specifically true* for the right-handers being *not true* for the left-handers, and symmetrically.

To this end, we designed the following new scheme: First, we construct the basis B_{RL} of implications which are holding for the two populations -joined by union into a $S_{84 \times A_{11}}$ matrix-, and which represents a *consensus* of what can be inferred from them. Then, we construct the list of all implications which have to be added to the

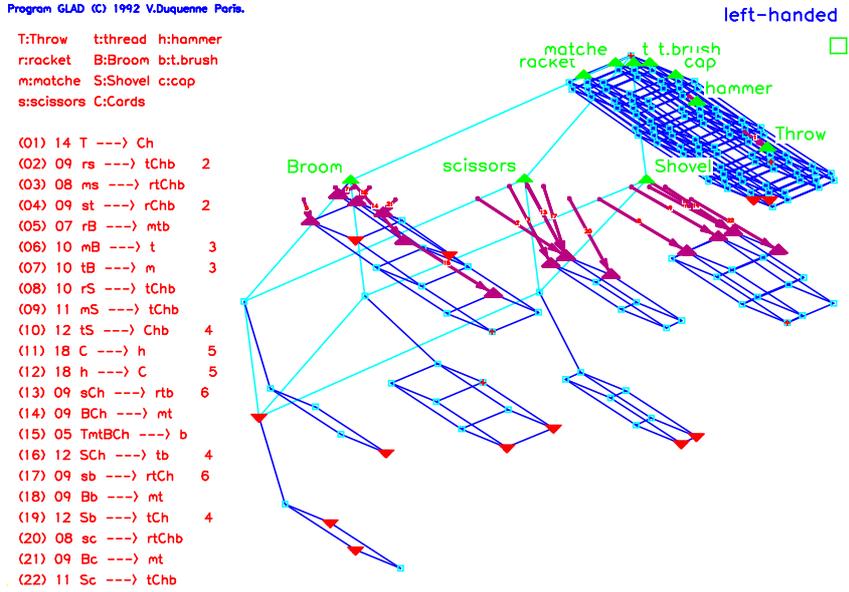


Fig. 3. Galois lattice and implications (left-handers)

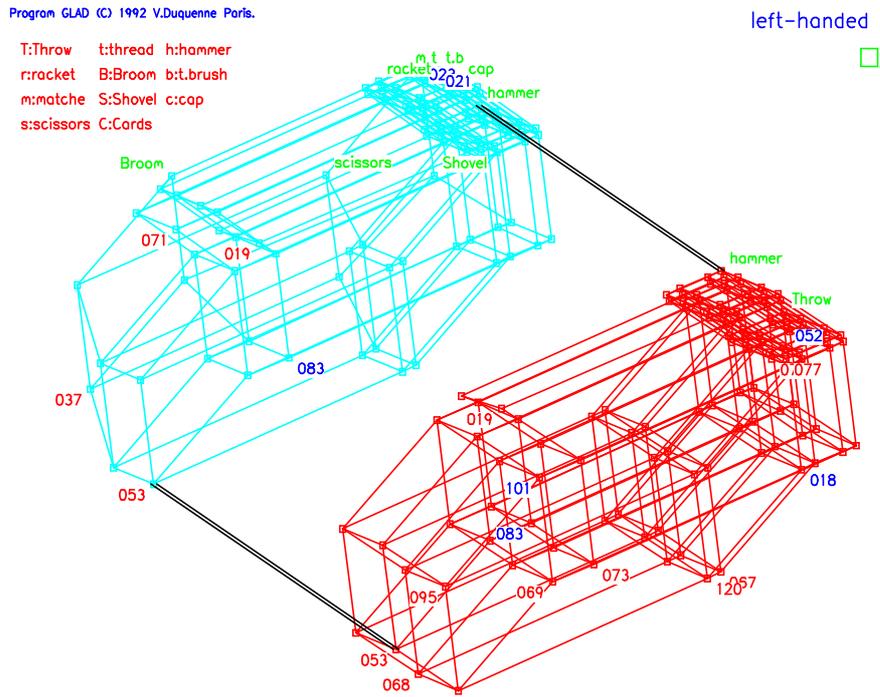


Fig. 4. Gluing decomposition in intervals (left-handers)

14 rmBc ---> b	14 TBhc ---> C	62 C ---> h	16 mBS ---> hb
16 rmtB ---> b	15 rBhc ---> C	25 TB ---> h	
25 mtSh ---> b	38 Tmhc ---> C	28 Sc ---> h	27 Bb ---> m
29 TmSh ---> b	40 rmhc ---> C	29 tS ---> h	
14 TmtBh ---> b	18 TtShc ---> C	44 Tc ---> h	57 Ts ---> r
27 rmsSh ---> b	19 rtShc ---> C	50 Tt ---> h	58 sh ---> r
12 TmBChc ---> b	33 Tthbc ---> C	56 Tb ---> h	
34 rmsChc ---> b	34 rthbc ---> C	14 rBS ---> h	40 sc ---> rh
		23 rsB ---> h	47 st ---> rh
		33 rSb ---> h	21 msB ---> rh
		51 rsb ---> h	

Basis B_{RL} (84 subjects)

10 mtBh ---> b
23 mthc ---> b
09 mBhc ---> b
39 rt ---> h
33 mt ---> h
21 B ---> h
33 tb ---> h
33 rc ---> h
27 tc ---> h
30 bc ---> h
51 T ---> r
44 Ch ---> r
38 Trth ---> s
32 Trhc ---> s
50 rs ---> T
20 rBh ---> T
43 rb ---> Th
32 rmth ---> Ts
26 rthc ---> Ts

Basis B_{R-RL} (61 subjects)

T:Throw
r:racket
m:matches
s:scissors
t:thread
B:Broom
S:Shovel
C:Cards
h:hammer
b:t.brush
c:cap

Actions

18 h ---> C
14 T ---> Ch
10 tB ---> m
09 BCh ---> mt
09 Bc ---> mt
07 rB ---> mtb
09 sb ---> rtCh
08 ms ---> rtChb
10 mB ---> t
12 SCh ---> tb
12 Sb ---> tCh
09 rs ---> tChb
10 rS ---> tChb
11 mS ---> tChb

Basis B_{L-RL} (23 subjects)

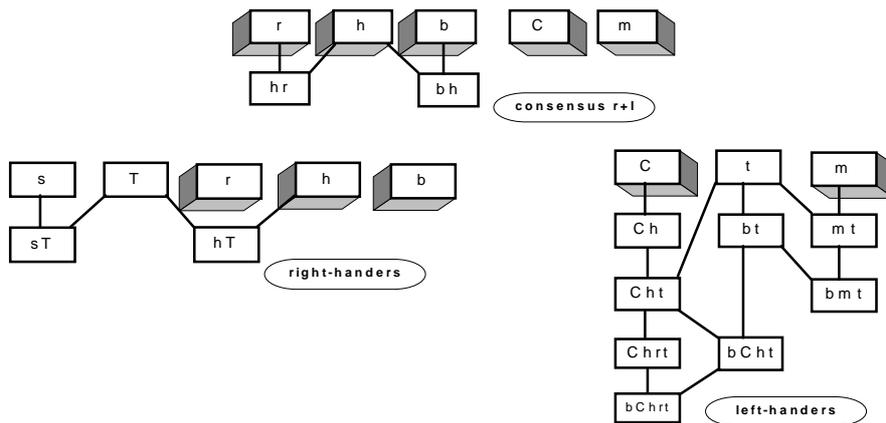


Fig. 5. The implications that are common / specific to right / left-handers

basis B_{RL} in order to generate a list being equivalent to the basis B_R of the right-handers, which can be summarized by a new *specific basis* that will be denoted by B_{R-RL} . Symmetrically, for left-handers, the specific basis B_{L-RL} is defined. Comparing and commenting on the consensus basis and these two specific basis should reveal what is common to both and what is specific to each of these populations, which addresses directly the cognitive structure of left / right-handedness.

For the results concerning the consensus by union of right / left-handed subjects, the 35 implications of the common basis B_{RL} have only seven distinct conclusions, which can be organized in two groups along an axis, from the actions that require more energy, *r:racket*, *h:hammer* and *b:tooth-brush* (and *hr* and *bh*, see Fig.5 top), up to those involving dexterity, such as *C:Cards* and *m:matches*.

Specific properties of the right(/ left)-handed subjects: The 19 implications of the specific basis B_{R-RL} are also structured in two groups. The first twelve ones refine the previous group of the consensus concerning energy, and imply either *r:racket*, *h:hammer* or *b:tooth-brush*, while the second group implies new conclusions: *T:Throw* and *s:scissors*. Hence, most of the implications for the right-handers imply an action which requires more energy than dexterity, to the exception of *s:scissors* which plays a peculiar role. As it is the case for the consensus by union, the partial order of conclusions is simple, being of order dimension five and length two.

Specific properties of the left(/ right)-handed subjects: As opposed to the right-handers, all the conclusions of the specific basis B_{L-RL} involve more dexterity, with *C:Cards* and *m:matches*, as in the consensus, but also the new conclusion *t:thread*, in conjunction with other actions, up to generate a complex partial order of conclusions of dimension three, and length five (see Fig.5 bottom).

Hence, to temporarily conclude with this study, as the items of the laterality questionnaire do not have the same categorical impact for left / right-handers, it can be stressed that they do not generate the same structures, which confirms and extends previous data from literature. The behavior of right-handers seems globally more stereotyped, with a minority of subjects generating a rich set of dependencies between conjunctions of actions. The behavior of left-handers appears less stereotyped and governed more by independence and avoidance among actions. While the implications holding in both populations are clearly scaled on an axis energy / dexterity, it is significant that right-handers refine the conclusions governed by energy, while left-handers those involving dexterity. Now, when some questionnaires pretend to evaluate laterality along a *continuum*, this analysis seems to question such a strong assumption. This calls for cooperation to investigate other populations of subjects, to be in a position to refine and to assess the cognitive structure of handedness.

3 Implications between two sets of variables: partial trisomies 21

A main question in Behavioral Genetics concerns the assessment of relationships between genotypic / phenotypic variables. We now report how these questions can be revisited through *lattice analysis* in the case of Down syndrome (see [4] for more interpretations). Its phenotype and genotype involve morphologic and anatomic abnormalities with more or less severe mental retardation, and a partial or complete

triplication of chromosome 21. This study will extend a previous work (see [7]) on the "Molecular mapping of twenty-four features of Down syndrome on chromosome 21" that showed that a candidate region for Down syndrome ("DCR") is the band q22.2.

Matching model. After an analysis of the genotypes (11 cytogenetic *bands*) and phenotypes (24 *features*) of a population of ten patients suffering from partial trisomy, which can be coded into a binary table $S_{10 \times (B_{11} + F_{24})}$, the matching was defined there by a set theoretic procedure: each feature $f \in F_{24}$ was ascribed to the subset of bands -say $f^{FB} \subseteq B_{11}$ - that are shared by all patients "having f ", which can be read as an *implication* $f \rightarrow f^{FB}$ (and not a *correlation* as sometimes said). The *molecular mapping* onto *minimal regions* is a representation of these implications in terms of their locations on chromosome 21.

Our discussion of the original model starts with a basic question: which are the respective interests and genetic significance of the implications "features \rightarrow bands" and "bands \rightarrow features"? Both kinds are equally informative -and cannot be ignored- as they address the two sides of *genetic causation of observed phenotypes*.

The latter point out to the common features of a group of patients which are sharing a set of bands, which expresses one -out of several possible- *sufficient condition for having these common features* -sufficient, since other patients, while not sharing these bands, may also have these features, due to "*some other causes*".

Conversely, an implication "feature \rightarrow bands": $f \rightarrow f^{FB}$ expresses a -unique, in contrast- maximal *necessary genotypic condition for having this phenotype* -necessary, since all patients having it do share these bands in common without exception, while other patients may also share these bands without having this phenotype due to "*some variability in the expression of responsible genes*". This variability could also be ascribed by assigning a *penetrance* to different genetic configurations. Even mental impairment, the only constant finding, varies in its expression and severity. Hence, the viewpoint taken in the original papers (see [7,19]) stresses the "feature \rightarrow bands" implications as formalizing a set theoretic *matching*, hence assuming a *genotypic / phenotypic asymmetry* and focusing only on the unique maximal *necessary genotypic condition for having each single features* $f \in F_{24}$.

In this report, we reinvestigate the model, in order to evaluate how far these implications "feature \rightarrow bands": ($f \rightarrow f^{FB}$ for all $f \in F$) are from defining equivalencies $f \leftrightarrow f^{FB}$ -in which case all the patients sharing f^{FB} do have f as well, so that f^{FB} is of maximal *penetrance*-, and this question can be either raised locally, or globally. Locally: such an equivalence is of greater confidence in view of further gene identification since it expresses a *necessary and sufficient* condition for having feature f , hence providing a *genotypic characterization* of it without intervention of "other causes" and "expression variability": all possible minimal sufficient conditions are then confused with the maximal necessary one in a unique region. Globally: if all these equivalencies held, that would assess that the set of bands *is sufficient* to characterize the polymorphism of the *phenotypic diversity* of trisomy 21 in this population. It turns out that only five of the 24 features bear such an equivalence, which partly explains the original genotype / phenotype asymmetry assumption that can be found in literature (see [7,19]), and justifies to extend the model by relaxing / extending the assumptions to refine the evaluation.

Extended model. This evaluation requires some steps: we re-defined the *matching model* by introducing appropriate denotations to study the *redundancy* of these implications $f \rightarrow f^{FB}$ regarding the phenotypic structure of the population -whenever they appear to be the consequence of some implications between features or differ greatly in their extensions in the population, being therefore far from equivalencies.

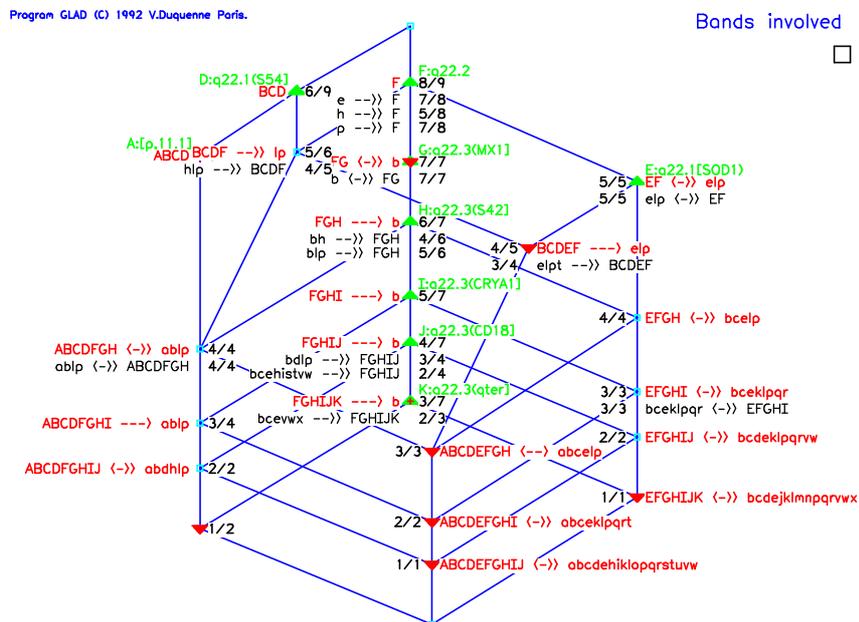
This more generally leads to consider implications / equivalencies *between conjunctions* of bands *as well as* features, and therefore to evaluate the complexity of the genotypic / phenotypic structures by using the *Galois Lattices* associated to the binary tables $S_{10 \times B_{11}} / S_{10 \times F_{24}}$. Extending the model up to conjunctions involves a risk of "combinatorial explosion" which is tamed by the Galois lattice construction that focus only on *pertinent* (i.e. maximal shared by some group of patients) *conjunctions* and has two major benefits. First, this treats the syndromes *as they are*: conjunctions of features shared by some group of patients, and this increases the chance of detecting genotypic characterizations through equivalencies, since feature conjunctions obviously have smaller extensions than their single features.

It should also be noted that lattice analysis confirms and makes more precise a previous conclusion that DCR is structured around band q22.2 (see [7]), but unfolds it along three independent directions q22.1(S54) (up to S54), q22.1[SOD1) (from SOD1), and a chain of nested intervals of q22.3 (see Fig.6 top).

The genotype / phenotype lattices are labeled with *basis* of implications going *from feature conjunctions into band conjunctions*, and conversely, which *completely characterizes the matching* of the geno / phenotypic structures in terms of necessary / sufficient conditions for having the phenotypes. It is shown that half of the band conjunctions which characterize the genotypic structure bear equivalencies -hence of maximal penetrance-, which define a *core structure* "already assessed" (see Fig.6).

Hence, the cytogenetic bands B_{11} actually provide a coherent core structure, even if they do not completely characterize the phenotypes yet -which would require more patients or a finer description of their genotypes. This can be interpreted as an encouraging result that also gives some strength to the matching model's extension, which both treats the syndromes as conjunctions of features and takes in charge the potential polymorphism of trisomy 21. Notice that the *extended model* departs only from the *original* one by: (1) *assuming* a more symmetric view on the genotypes / phenotypes and necessary / sufficient genotypic conditions, (2) *weighting* the implication premises / conclusions for evaluating their discrepancy to equivalence, and (3) *extending* the matching procedure to the conjunctions of features.

The resulting combinatorial complexity is minimized by the *Galois lattices* which focus only on *pertinent conjunctions*, while this extension is essential to reveal the phenotypes / genotypes for which there is actually a genotypic characterization of the phenotypes -or not, suggesting some further refinements-, for the cytogenetic bands under study. This should lead to raise local questions for bettering the genetic description of features, and to search for some minimal sets of questions on conjunctions -not bearing equivalencies- of middle size extensions -i.e. reasonably high in the lattices-, which should be investigated through considering *new smaller chromosomic regions*, and to extend the study by constructing *larger data bases*.



ABCDEFGHIJK					
0111000000			BCD	---	6/9
#ML 1111000000			ABCD	---	5/9
0000010000	7, 5, 7/8	e, h, p	---	---	8/9
0111010000	4/5	hlp	---	BCDF	5/6
0000110000	5/5	elp	<-->	EF	5/5
#FG 0111110000	3/4	elpt	---	BCDEF	4/5
#IG 0000011000	7/7	b	<-->	FG	7/7
0000011100	4, 5/6	bh, blp	---	FGH	6/7
1111011100	4/4	ablp	<-->	ABCDFGH	4/4
0000111100			<-->	EFGH	4/4
#TY 1111111100			<-->	ABCEFGH	3/3
0000111100			---	FGHI	5/7
1111011100			---	ABCEFGHI	3/4
0000111100	3/3	bceklpqr	<-->	EFGHI	3/3
#LI 1111111100			<-->	ABCEFGHI	2/2
0000011110	2, 3/4	bcehistvw, bdlp	---	FGHIJ	4/7
1111011110			<-->	ABCEFGHIJ	2/2
0000111110			<-->	EFGHIJ	2/2
#AL 1111111110			<-->	ABCEFGHIJ	1/1
#AB 0000011111	2/3	bcevw	---	FGHIJK	3/7
#SC 1111011111			---	ABCEFGHIJK	1/2
#DL 0000111111			<-->	EFGHIJK	1/1
ABCDEFGHIJK					
. K:q22.3(qter)					
. J:q22.3(CD18)					
. I:q22.3(CRYA1)					
. H:q22.3(s42)					
. G:q22.3(MX1)					
. F:q22.2					
. E:q22.1[SOD1]					
. D:q22.1(S54)					
. C:q21					
. B:q11.2					
. A:[p,11.1]					

Fig. 6. Top: the band lattice is labeled with all implications "bands → features", and non-redundant implications "features → bands". Bottom: the band lattice is listed as an extended molecular mapping with the subject bands and implications.

4 Conclusion

On a generic level, these two studies show that *Galois lattices* and *implication basis* are most useful for exploring the matching of two binary relations (0/1 data tables), which are paired either "by the columns" or "by the rows". The former situation leads to the fundamental problem of *comparing* two Galois lattices -or closure operators / implication basis- by taking a "differential" viewpoint. Dually, the latter case leads to search what can be inferred, in the same population, *from* a set of variables *to* another one, with an extra bonus for the characterization of equivalencies that are indeed approaching "causality". On an interpretation level, both studies call for exploring larger populations for refining the provisory conclusions.

This raises two kinds of questions, which are addressed either to the *psychologist / biologist* in charge of the data and its interpretations, or to the *computer scientist* who would like to extend these approaches to much larger data bases in the successful topic of *KDD (Knowledge Discovery in Data bases)*, or to both.

First, the underlying structures are here *simpler* and therefore *more natural* than what is commonly used in these disciplines: extra methodological assumptions cost nothing, for sure, but ... plausibility. To the computer scientist a first remark: our experience is that scanning thousands of lattice elements / implications cost nowadays no more than one second with reasonable programs and pc-computers. Hence, the current question is perhaps more to do something about it, to promote and unfold *semantic interpretations*, than to cut this time by a quarter: it is a question of priority. Another wisdom message will be that whenever one adds a lot more subjects (rows, information), all observed lattices become Boolean -without implications!- that stresses an urge for a real, funded, sensitive *approximation theory* of these discrete structures. Facing this apparent contradiction between practical / theoretical needs, all the announced developments for pruning, navigating ... or putting more flexibility through this combinatorial complexity will be welcome: the remaining questions will of course be their canonicness and significance, their efficiency and usefulness.

Acknowledgments

This research was supported by the Action Bioinformatique 2000, Centre National de la Recherche Scientifique - Centre Auvergne Limousin, Conseil de la Région Centre and the Fondation Jérôme Lejeune. The 2nd and 5th authors received a grant from the Fondation pour la Recherche Médicale. The development of the program GLAD was aided by CNRS, University R. Descartes, and Maison des Sciences de l'Homme-Paris.

References

1. Birkhoff G.: Lattice Theory. 3rd. edn. AMS, Providence (1967).
2. Burosh G., Demotrovits J., Katona G.O.H.: The poset of closures as a model of changing data bases. Order 4 (1987) 127-142.

3. Caspard N., Monjardet B.: The lattice of Moore families and closure operators on a finite set. To appear in *Discrete Mathematics* (2000).
4. Chabert C., Cherfouh A., Delabar J.-M., Duquenne V.: Assessing implications between genotypic / phenotypic variables through lattice analysis. To appear in *Behavior Genetics* (2001).
5. Crapo H.: Selectors: a theory of formal languages, semimodular lattices, and branching and shelling processes. *Advances in Math.* 54 (1984) 233-277.
6. Day A.: The lattice theory of functional dependencies and normal decompositions. *Intern. J. of Algebra and Computation* 2 (4) (1992) 409-431.
7. Delabar J.M., Theophile D., Rahamani Z., Chettouh Z., Blouin J.L., Prieur M., Noel B., Sinet P.M.: Molecular mapping of twenty-four features of Down Syndrome on chromosome 21. *Europ. J. of Human Genetics* 1 (1993) 114-124.
8. Doyen A.-L., Duquenne V., Nuques S., Carlier M.: What can be learned from a lattice analysis of a laterality questionnaire? To appear in *Behavior Genetics* (2001).
9. Duquenne V.: What can lattices do for experimental designs? *Mathematical Social Sciences* 11 (1986) 243-281 (MR #87j:62140).
10. Duquenne V.: On the core of finite lattices. *Discrete Math.* 88 (1991) 133-147 (MR #92g:06011).
11. Duquenne V.: Towards an intensional logic of symptoms. *Current Psychology of Cognition* 15 (3) (1996) 323-345.
12. Duquenne V.: Latticial structures in Data Analysis. *Theoretical Computer Science* 217 (1999) 407-436.
13. Duquenne V.: Le Pape A., Moad M., Temam K.: Representing Two-mode Binary Data as Galois Lattices, JMS 1999 Baltimore abstract #563, to appear in *Proc. Amer. Stat. Assoc.* (2000).
14. Ganter B.: Two basic algorithms in concept analysis, Preprint THD Darmstadt n° 831. German transl. in B. Ganter, R. Wille, K.E. Wolf (eds): *Beitrage zur Begriffsanalyse*, B. I. Wissenschaftsverlag, Mannheim (1987) 241-254.
15. Ganter B., Wille R.: *Formal concept analysis: mathematical foundations*. Springer, Berlin (1999).
16. Godin R., Missaoui R., Alaoui H.: Learning algorithms using a Galois lattice structure. *Proc. of the IEEE Intern. Conf. on Tools for Artificial Intelligence*, San Jose (1991) 22-29.
17. Guénoche A.: Construction du treillis de Galois d'une relation binaire. *Math. Info. Sci. Hum.* 109 (1990) 41-53.
18. Guigues J.-L., Duquenne V.: Familles minimales d'implications informatives résultant d'un tableau de données binaires. *Math. et Sci. Hum.* 95 (1986) 5-18 (MR #87k:08009).
19. Korenberg J. R. *et al.*: Down syndrome phenotypes: The consequences of chromosomal imbalance. *Proc. Natl. Acad. Sci. USA* 91 (1994) 4997-5001.
20. Speed T.P.: What is analysis of variance? *The Annals of Statistics* 15 (3) (1987) 885-941.
21. Torjman S. *et al.*: Identification of clinical subgroups of infantile autism: A preliminary step for a genetic study. *American J. of Medical Genetics* 74 (1997) 572.
22. Wild M.: Implicational bases for finite closure systems. In W. Lex (ed.): *Arbeitstagung Begriffsanalyse und Künstliche Intelligenz*. TU Clausthal (1988) 147-169.
23. Wild M.: A theory of finite closure spaces based on implications. Preprint MIT (1991), Cambridge (MA). Published in *Advances in Math.* 108 (1994) 118-139.
24. Wille R.: Knowledge acquisition by methods of formal concept analysis. In E. Diday (ed.): *Data analysis, learning symbolic and numeric knowledge*. Nova Sci., New York (1989) 365-380.