ANALYSES DE SIMILARITES ET D'IMPLICATIONS ENTRE PROCEDURES D'ELEVES DANS DE COURTES DEMONSTRATIONS DE GEOMETRIE

Annie LARHER
Lycée Ile-de-France
et Equipe de Didactique de l'IRMAR
Rennes

I. Introduction

C'est principalement en géométrie que les élèves mettent en œuvre le raisonnement déductif. Les difficultés rencontrées sont très importantes : elles hypothèquent quelquefois la suite de leur scolarité mathématique. L'étude présentée ici vise à connaître, dans certaines situations, l'origine et la nature des erreurs les plus fréquentes, à analyser les procédures utilisées par des élèves de 1er cycle dans un contexte de démonstration en géométrie, en particulier dans le cas où l'activité déductive se réduit à une simple inférence.

Cette analyse ne peut manquer d'utilité puisque l'inférence simple est un passage obligé pour toute démonstration élémentaire exigée des élèves au début de leur apprentissage de la preuve mathématique. Or il est bien difficile, voire impossible, pour l'enseignant, de repérer à chaque fois dans une copie d'élève le type d'erreur commise et surtout sa répétition, sa fréquence dans la classe et les conditions dans lesquelles l'erreur s'élabore et apparaît. Il lui est encore plus difficile de trouver pour chaque élève les situations qui permettraient de déstabiliser et mieux, d'éliminer les procédures erronées.

L'ordinateur, en revanche, permet un travail plus individualisé ; en outre, sous certaines conditions, il révèle les démarches des élèves et facilite aussi leur analyse didactique. Ainsi le logiciel «Premier Pas», élaboré par A. Simon*, remplit ces différentes fonctions : il renforce l'apprentissage du fonctionnement d'un pas déductif, permet le bilan et le recensement des acquis et, surtout, fournit des indications plus précises et objectives que les travaux papier-crayon sur les comportements des élèves placés dans des situations élémentaires.

Les méthodes statistiques classiques ne permettant pas d'accepter ou de réfuter une hypothèse didactique, nous utilisons ou construisons des méthodes multidimensionnelles d'analyses de données qui, de leur côté, dégagent de grandes structures de comportements erronés. Ce sont :

- la classification hiérarchique, symétrique, permettant d'étudier les proximités, les similarités entre les erreurs,

* A. Simon est membre de l'équipe de didactique de l'Institut de Recherche Mathématique de Rennes (IRMAR).

«petit x» n° 32 pp. 25 à 39, 1992-1993

- et la classification implicative entre variables binaires qui donne un sens statistique à la relation dissymétrique :

«si on observe telle erreur x, alors on observe aussi telle erreur y».

L'analyse implicative, prolongée à d'autres types de variables et à des classes de variables (ici, de procédures) munit l'ensemble de celles-ci d'une dynamique orientée et permet, par regroupement des procédures en classes, de donner un sens à l'énoncé plus large :

«si telle classe d'erreurs A apparaît, alors telle classe d'erreurs B apparaît également ».

Cette extension permet, par la construction d'un arbre de classes orientées, de dégager de grandes familles de comportements ou de conceptions erronés des élèves. Elle apporte ainsi des éléments de réponse à quelques questions levées par notre problématique actuelle en didactique des mathématiques :

- à un niveau de cursus donné, peut-on, dans une situation donnée, déterminer une hiérarchie partiellement ordonnée de procédures de résolution de problèmes de mathématiques, signe d'une connaissance en voie de constitution ?
- à un niveau de cursus donné, peut-on définir, à partir de classes ordonnées de procédures, des conceptions homogènes et résistantes relativement à un certain savoir ?

Nous montrons ici les principaux résultats obtenus sur le plan didactique en faisant fonctionner concurremment les différents outils évoqués ci-dessus - et en particulier l'implication statistique entre classes - sur les données recueillies à deux questionnaires de géométrie proposés à des élèves de 1er cycle en apprentissage du raisonnement déductif. Nous ne prétendons pas que cet apprentissage doive se réduire à la résolution d'exercices consistant, comme ceux de ces questionnaires, à compléter des inférences. Nous ne prétendons pas non plus que le type d'exercice proposé ici contribue à donner du sens à l'activité de preuve mathématique. Il s'agit seulement de donner les principales hypothèses que l'analyse implicative entre classes et les autres analyses nous ont amenés à formuler sur des comportements ou des conceptions d'élèves placés dans des situations élémentaires de démonstration déductive en géométrie.

Mais revenons tout d'abord sur les méthodologies de révélation et d'analyse des démarches des élèves que nous avons utilisées.

II. Méthodologie retenue

1. Méthodologie de révélation

a) Le logiciel « Premier Pas »

Ce logiciel propose à l'élève une liste de faits mathématiques (géométriques en l'occurrence) pouvant tenir lieu, suivant les situations, d'hypothèses ou de conclusions ainsi qu'une liste de théorèmes. Faits et théorèmes sont repérés par des numéros. Une inférence incomplète étant proposée, l'élève doit choisir un ou plusieurs faits, un ou plusieurs théorèmes pour que soit validée l'inférence (ou les inférences successives).

Exemple

(AB) // (EF) et (CD // (EF)	?	(AB) // (CD)
Hypothèse	Théorème	Conclusion

Nécessairement, la stratégie de décision de l'élève dans cet exercice très particulier est fort proche de celle déployée dans les Q.C.M. et en revanche, très différente de celle qui est suivie dans les démonstrations à plusieurs pas et dans les problèmes ouverts. Ici, l'élève n'a pas de véritable activité créatrice et le sens global d'un problème n'est pas mobilisable ; les seuls points d'appui sont le sens du pas de démonstration et l'ensemble langagier des assertions ou théorèmes dont il dispose.

Nous avons cependant remarqué, grâce à la répétition, à l'accumulation et à la concomitance d'erreurs, la stabilité de certaines procédures qui correspondent à des modèles de fonctionnement en équilibre aussi bien chez un élève particulier que chez l'élève en général. Nous en voyons des exemples en analysant les résultats obtenus à deux questionnaires proposés sur le logiciel «Premier pas» à des élèves de 5ème après l'enseignement de quelques propriétés de la symétrie par rapport à un point.

b) Présentation des questionnaires

Les questionnaires, appelés «6 Questions» et «5 Questions» que nous étudions ici se rapportent donc à la symétrie centrale et, pour une question du 1er (Q_5) à la transitivité du parallélisme. Les questions posées visent à étudier l'effet d'un certain nombre de variables liées à la logique de l'inférence et à la forme des énoncés proposés.

Hypothèses et théorèmes sont donnés : l'élève doit compléter en choisissant un des onze faits de la liste à titre de conclusion. A la suite d'un premier essai, il est autorisé à faire un deuxième et dernier essai. Les questions sont indépendantes.

Faits

1. (EF) et (CD) sont symétriques par rapport au point I
2. [MN] est le symétrique de [PR] par rapport au point I
3. (AB) et (CD) sont symétriques par rapport au point O
4. (MN) // (PR)
5. (CD) // (EF)
6. (AB) // (CD)
7. (AB) // (EF)
8. MN = PR
9. CD = EF
10. AB = CD
11. AB = EF

Théorèmes

1. La symétrie centrale conserve les longueurs.
2. Si (D) // (D') et (D') // (D"), alors (D) // (D").
3. Le symétrique d'une droite (D) par rapport à un point est une droite (D') parallèle à (D).
4. Si deux droites sont symétriques par rapport à un point, alors elles sont parallèles.
5. Deux segments symétriques par rapport à un point ont même longueur.
6. La symétrie centrale conserve les directions.

Chaque question se présente schématiquement ainsi :

Hypothèse : fait n° a - Théorème n° b - conclusion : fait n° ?

Voici l'ensemble questions-réponses du questionnaire «6 Questions».

	Hypothèses	Théorèmes	conclusions à trouver
Q_1 {Hypothèse (H) : 1, Théorème (T) : 3, Conclusion (C) : 5	(EF) et (CD)symétriques par rapport à I.	Le symétrique de (D) par rapport à un point est (D') // (D).	(EF) // (CD)
Q_2 {H : 3, T : 4, C : 6	(AB) et (CD) symétriques par rapport à O.	Si 2 droites sont symétriques par rapport à un point, alors elles sont parallèles.	(AB) // (CD)
Q_3 {H : 2, T : 5, C : 8	[MN] est symétrique de [PR] par rapport à I.	2 segments symétriques par rapport à un point ont même longueur.	MN = PR
Q_4 {H : 3, T : 6, C : 6	(AB) et (CD) symétriques par rapport à O.	La symétrie centrale conserve les directions.	(AB) // (CD)
Q_5 {H : 6 et 5, T : 2, C : 7	(AB) // (CD) et (CD) // (EF).	Si (D) // (D') et (D') // (D") alors (D) // (D").	(AB) // (EF)
Q_6 {H : 2, T : 1, C : 8	[MN] est symétrique de [PR] par rapport à I.	La symétrie centrale conserve les longueurs.	MN = PR

Le questionnaire «5 Questions» soumis aux élèves postérieurement au «6 Questions» pour conforter des hypothèses au sujet des démarches de démonstration, ne diffère de celui-ci que :

- par la suppression de la question Q_5 non relative à la symétrie ponctuelle (il est apparu en effet que les réponses au «6 Questions» étaient un peu biaisées par la présence de cette question relevant d'un autre concept) ;

et - par l'ajout de deux faits dont l'absence dans le «6 Questions» semble avoir provoqué des confusions trop singulières :

12. (AB) = (CD)
13. [MN] = [PR].

Chaque modalité de réponse est codée par un triplet. Exemple : 3-6-10 (Q_4).

2. Méthodologie d'analyse

Relativement aux 80 élèves ayant répondu au questionnaire «6 Questions», nous avons relevé en tout 31 modalités de réponse à l'ensemble des 6 exercices (premier et second essais).

Voici le tableau de ces 31 modalités et de leurs nombres d'occurrences :

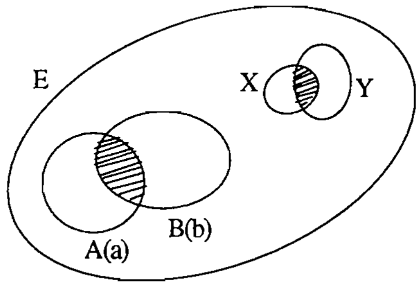
Questions	Modalités	N°	Nombre d'occurrences
Q_1	1-3-5 (R_1)	1	63
	1-3-1	2	14
	1-3-2	3	6
	1-3-3	4	10
	1-3-6	5	4
	1-3-9	6	9
Q_2	3-4-6 (R_2)	7	77
	3-4-2	8	4
	3-4-5	9	4
	3-4-3	10	2
	3-4-10	11	3
Q_3	2-5-8 (R_3)	12	70
	2-5-4	13	12
	2-5-6	14	4
	2-5-7	15	4
	2-5-2	16	4
Q_4	3-6-6 (R_4)	17	58
	3-6-3	18	17
	3-6-4	19	4
	3-6-5	20	6
	3-6-9	21	4
	3-6-10	22	41
Q_5	6,5-2-7 (R_5)	23	68
	6,5-2-3	24	7
	6,5-2-9	25	6
	6,5-2-11	26	11
Q_6	2-1-8 (R_6)	27	70
	2-1-3	28	12
	2-1-4	29	5
	2-1-5	30	6
	2-1-6	31	4

Certaines réponses erronées sont relativement fréquentes et pourtant aberrantes, du moins à première vue ; et on peut être tenté de penser, de façon superficielle, que ces élèves, ou bien n'ont rien compris, ou bien n'ont pas répondu sérieusement au questionnaire.

Or les analyses hiérarchiques des similarités d'une part, des implications mutuelles entre les procédures d'autre part, permettent de trouver quelques explications objectives aux réponses jugées incohérentes ou aléatoires.

a) Analyse hiérarchique des similarités entre variables binaires

Pour définir la similarité de 2 variables binaires a et b, I.C. Lerman (Université Rennes I) a construit un indice [Lerman 81] qui ne prend pas seulement en compte le nombre d'individus d'une population E possédant à la fois a et b, il situe plutôt la valeur de ce nombre par rapport à la valeur «attendue» du cardinal de l'intersection entre deux parties aléatoires X et Y de même cardinaux que A et B (ensembles des individus possédant respectivement les attributs a et b).



En termes intuitifs, les deux attributs a et b sont d'autant plus voisins que le nombre d'individus les possédant l'un et l'autre est invraisemblablement grand, dans une hypothèse d'indépendance entre X et Y, selon un modèle probabiliste.

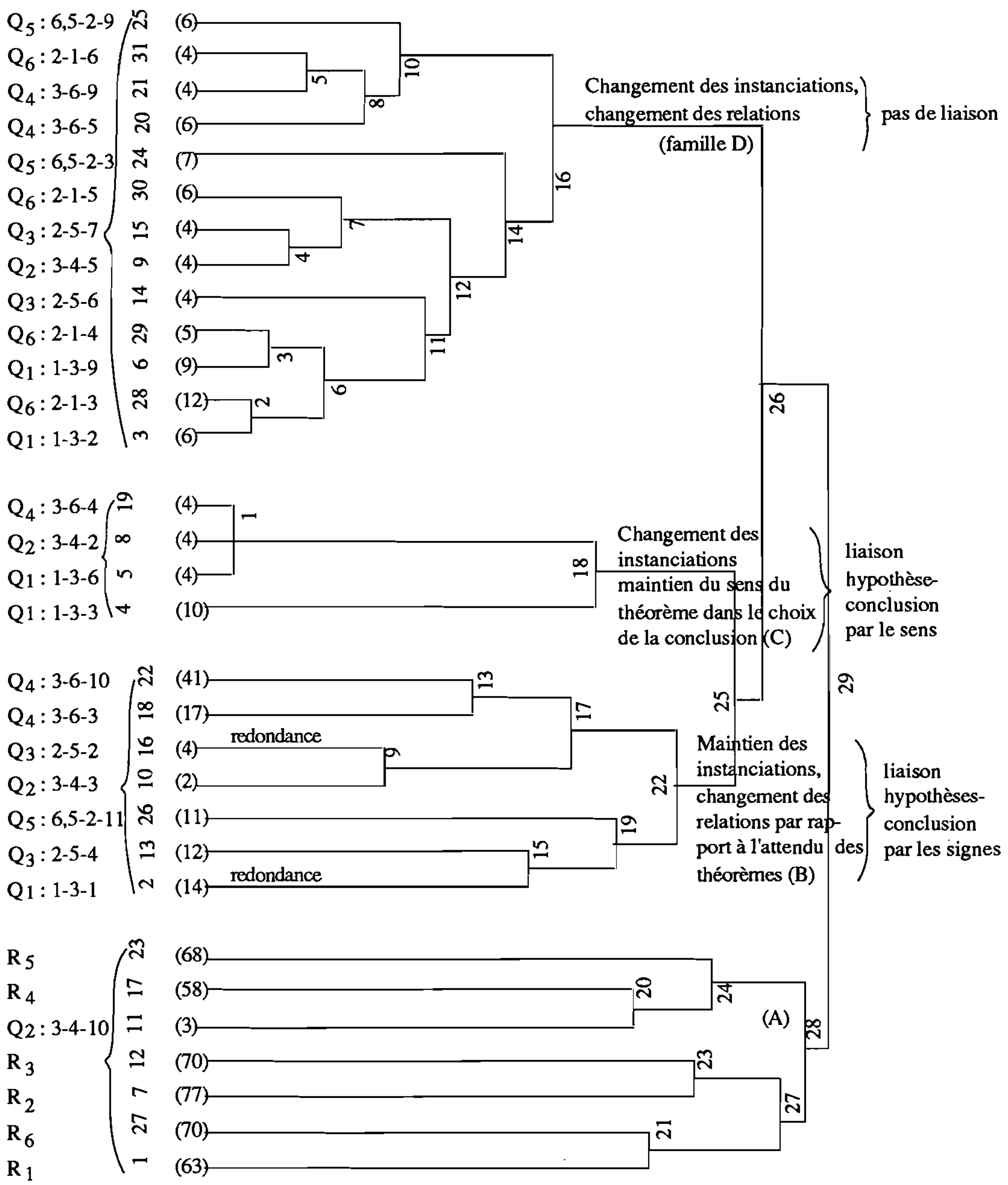
Dans le questionnaire «6 Questions», parmi les 31 modalités de réponses, quelles sont celles qui se «ressemblent» le plus ?

Les réponses sont rentrées sur ordinateur (logiciel C.H.I.C. de Saddo Ag Almouloud**) qui calcule tous les indices de similarité :

- entre les variables deux à deux : celles d'entre elles qui ont même indice de similarité se réunissent en une classe ;
- puis entre une variable et une classe ;
- enfin entre classes.

Les ressemblances ou différences entre les items ou les classes d'items sont traduites graphiquement sous la forme d'une hiérarchie.

** S. Ag Almouloud, équipe IRMAR - Rennes.



arbre des similarités 80 élèves

Quatre familles de modalités apparaissent très clairement :

A : toutes les bonnes réponses et la modalité 3-4-10 ;

B : procédures erronées, comportant toutes le maintien des noms des objets (segments, droites,...) mais le changement des relations par rapport à l'attendu des théorèmes (// remplacé par = ou = remplacé par //) ;

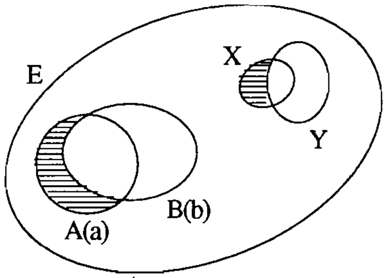
C : réponses d'élèves modifiant les instanciations (noms des objets) mais gardant le sens du théorème dans le choix de la conclusion ;

D : réponses incohérentes à la fois sur les variables et sur le sens du théorème.

b) Analyse implicative

• Intensité d'implication entre variables binaires

Pour définir les intensités d'implication mutuelle, R. Gras procède à partir de $A \cap \overline{B}$ comme I.C. Lerman à partir de $A \cap B$, selon un modèle probabiliste.
Soit Card A $\leq$ Card B. Si $A \subset B$, l'implication $a \Rightarrow b$ est vraie au sens logique dans E.
Si A n'est pas inclus dans B, mais si Card $(A \cap \overline{B})$ est relativement petit, on associe une mesure à $a \Rightarrow b$, dont la valeur est un indice de la qualité de la quasi implication de a vers b. C'est le cas par exemple où dans un questionnaire il y a peu d'élèves à avoir utilisé la procédure a sans avoir utilisé la procédure b.



En termes intuitifs, l'attribut a implique d'autant plus l'attribut b (B absorbe d'autant plus A) que le nombre d'individus possédant a et ne possédant pas b est invraisemblablement petit par rapport au nombre attendu, dans une hypothèse d'indépendance entre les parties aléatoires X et Y de mêmes cardinaux respectifs que A et B.

Un nombre, compris entre 0 et 1, appelé intensité d'implication, traduit de façon croissante la qualité de l'implication de a sur b.

Le graphe d'implication, autre originalité de la méthode, nous donne une représentation de l'ensemble des relations implicatives les plus fortes, opératoire pour l'analyse et l'interprétation. En particulier, à partir de ce graphe, nous émettons des conjectures du type : la réussite à une question donnée s'accompagne-t-elle de la réussite à une autre question ? Telle erreur implique-t-elle telle autre erreur ?

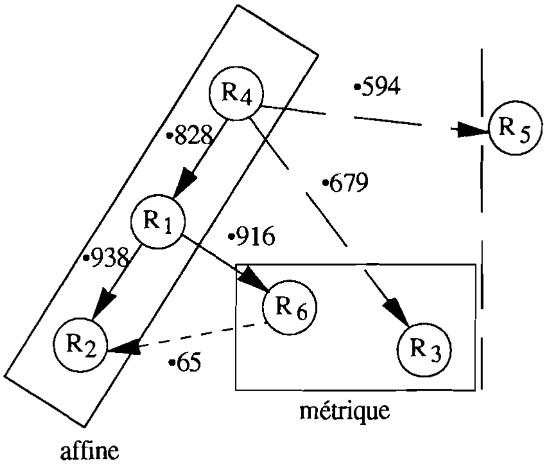
L'ordinateur donne le listing des intensités d'implication dont les valeurs gratifient de nuance les hypothèses formulées. Indiquons que le calcul de l'intensité d'implication peut être extrapolé à d'autres types de variables que les variables binaires : les variables modales et les variables quantitatives.

Voici le graphe implicatif des 6 réussites et les intensités d'implication pour le questionnaire «6 Questions» (R_i : réussite à la question (Q_i)).

- la coprésence de deux mots «conserver» et «directions» («conserver» avec le sens que lui ont donné, dans les classes, les enseignants qui ont soumis leurs élèves aux questionnaires) accroît la difficulté de la question (Q_4) ce qui explique la place de R_4 en racine de l'arbre ;

- la réussite semble liée à la nature du concept en jeu : séparation entre les 5 réussites relatives à la symétrie et la réussite R_5 relative à la transitivité. Seule existerait une liaison implicative (faible) de R_4 vers R_5 due à des capacités supérieures interconceptuelles ou simplement logiques ;

- R_3 et R_6 sont les réussites aux deux questions concernant les longueurs (propriété métrique M de la symétrie centrale). On peut s'étonner d'une part qu'elles n'admettent qu'une faible liaison entre elles, d'autre part que R_6 se trouve liée aux modalités R_1, R_2 et R_4 ayant trait au parallélisme de deux droites symétriques par rapport à un point (propriété affine de la symétrie centrale).



Sans doute ne faut-il pas chercher à cela d'autre raison que la présence du mot «conserver» dans le théorème de (Q_6) : nous présumons en effet que ce mot est lu par la majorité des élèves comme «confondre», d'où la bonne réponse MN = PR donnée à (Q_6) mais par erreur ! En revanche, il semblerait que la réussite à (Q_3), consécutive peut-être à un échec au 1er essai, corresponde davantage à une bonne compréhension du sens du théorème donné sous forme affirmative et sans mots ou difficiles ou ambigus.

• Implication entre classes de variables

L'implication entre classes ne prend véritablement sons sens qu'à condition qu'à l'intérieur de chaque classe dont on examine la relation avec d'autres, existe une certaine «cohésion» entre les variables qui la constituent.

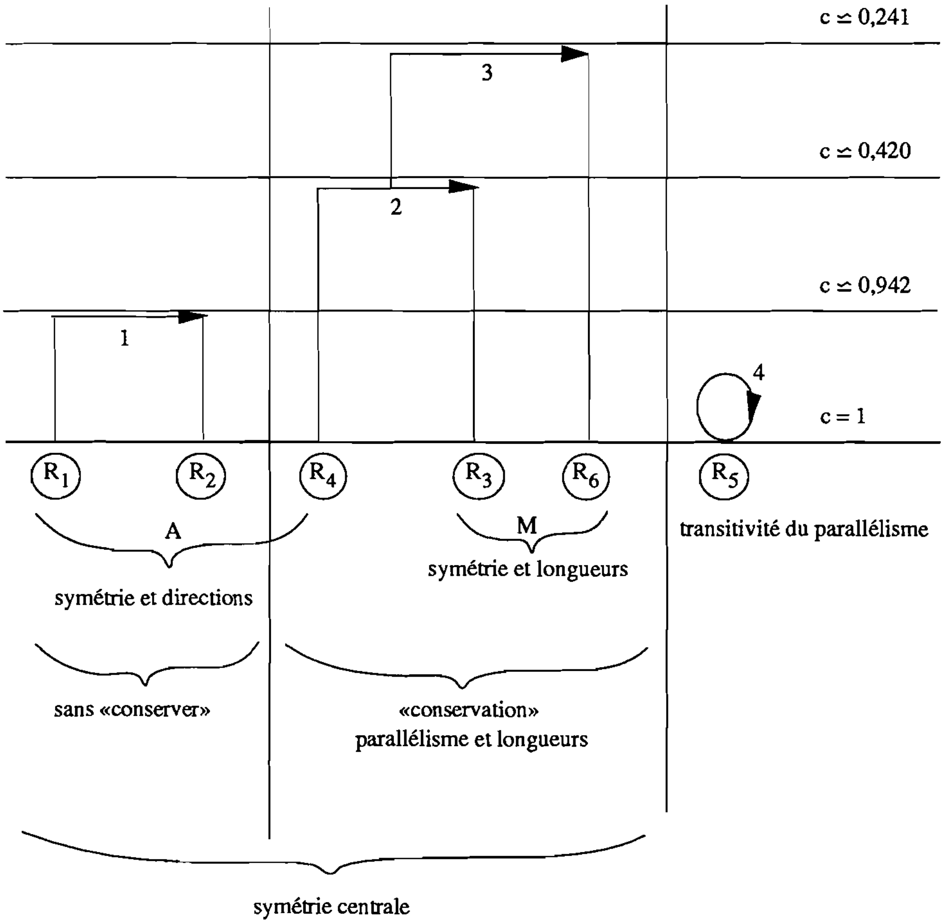
L'implication entre deux classes se constitue à partir des informations suivantes :
- les cohésions respectives des deux classes,
- une intensité d'implication extrémale des éléments d'une classe sur les éléments de l'autre,
- les cardinaux respectifs des deux classes.

Basé sur le concept d'entropie et s'y opposant, l'indice d'implication que nous avons retenu :
* croît avec les cohésions de chaque classe,
* croît avec la liaison extrêmale,
* décroît avec les cardinaux des classes.

Nous établissons un algorithme d'agrégations successives de classes où la cohésion des classes formées est toujours décroissante au sens large. Saddo Ag Almouloud [thèse 1992], a élaboré sur P.C. des programmes d'analyses de données incluant :

- le calcul d'intensités d'implication entre attributs ou variables binaires ou non,
- le calcul de cohésions et d'indices d'implication entre classes selon l'algorithme précédent,
- la construction de l'arbre d'implication entre classes sur le modèle algorithmique utilisé par R. Gras pour l'arbre de classification hiérarchique de I.C. Lerman.

Ces programmes ont servi aux traitements des modalités de réponse données par les élèves aux questionnaires «6 Questions» et «5 Questions» présentés ci-dessus ; par exemple, entre les réussites au «6 Questions», nous obtenons :



III. Analyses hiérarchique et implicative de l'ensemble des réponses aux questionnaires

L'examen des résultats bruts et l'analyse hiérarchique des similarités ont mis en évidence dans les questionnaires «6 et 5 questions» des types d'erreurs faites par de jeunes élèves en situation d'apprentissage de la démonstration mathématique. En fait, 3 grandes familles de procédures erronées, très stables mais non disjointes, se sont dégagées :

- **changement d'instanciation** : l'élève change le nom des objets,

- **répétition** : l'élève répète en conclusion l'hypothèse de l'inférence,

- **changement de relation** par rapport à l'attendu du théorème ou, plus spécifiquement, **confusion entre le parallélisme et l'égalité**.

Sur ces familles, l'application des nouveaux outils d'implication statistique a permis de donner des éléments de réponse aux questions de notre problématique, à savoir :

- telle erreur s'accompagne-t-elle d'une autre erreur plus fréquemment que d'une réussite ?
- dans le cas où un type d'erreur implique directement la réussite à une ou plusieurs questions : à quelle(s) question(s) ? dans quelle circonstance ?

Dans le graphique suivant, nous avons regroupé les modalités de réponse du «6 Questions» par familles de manière à mettre en évidence les implications de : réussites vers réussites, procédures erronées vers réussites, procédures erronées vers procédures erronées.

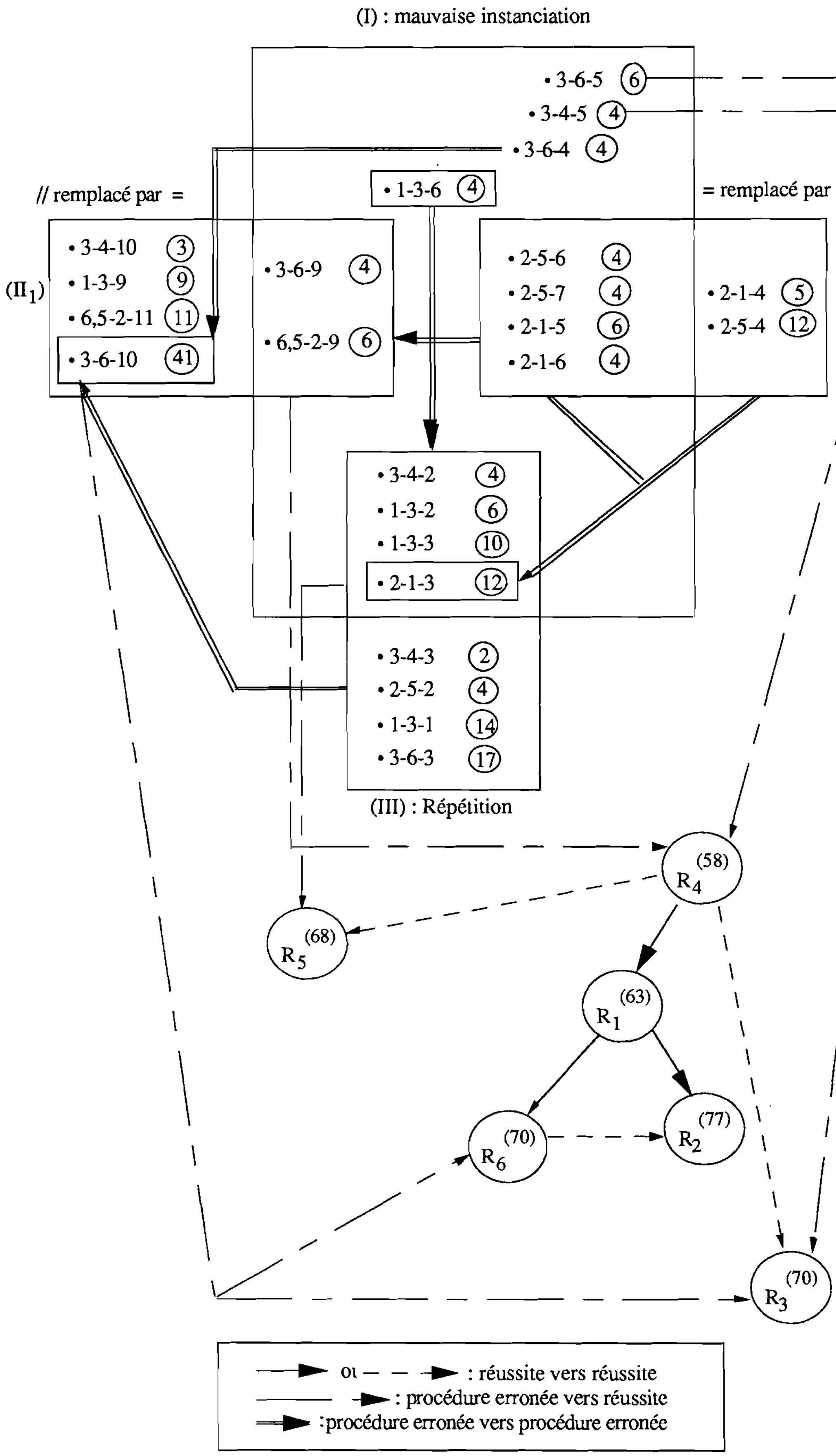
Rappelons que chaque modalité de réponse est codée par un triplet, par exemple, 3-4-3 signifie :

hypothèse : fait n° 3 ; théorème : n° 4 ; conclusion donnée : fait n° 3.

En clair :

Hypothèse : 3	Théorème : 4	Conclusion : 3
(AB) et (CD) sont symétriques par rapport au point O.	Si deux droites sont symétriques par rapport à un point, alors elles sont parallèles.	(AB) et (CD) sont symétriques par rapport au point O.

La dynamique orientée apparaissant sur le graphique implicatif ci-après entre les familles de procédures soulève naturellement des questions d'ordre didactique et cognitif. Des hypothèses peuvent être émises sur les différents types d'erreurs. Celles que nous formulons plus loin ont été mises à l'épreuve du 2ème questionnaire («5 Questions») et d'autres activités ; elles ont été confrontées à celles faites a priori, contrôlées par l'analyse hiérarchique des similarités ainsi que par une analyse factorielle de correspondances ; ces analyses corroborent, valident, enrichissent les résultats acquis par l'étude implicative tout en les spécifiant par rapport aux autres, de sorte que les interprétations s'appuient et se précisent mutuellement.



(I) : mauvaise instanciation
• 3-6-5 6
• 3-4-5 4
• 3-6-4 4
• 1-3-6 4
// remplacé par =
= remplacé par //
(II1)
• 3-4-10 3
• 1-3-9 9
• 6,5-2-11 11
• 3-6-10 41
• 3-6-9 4
• 6,5-2-9 6
• 2-5-6 4
• 2-5-7 4
• 2-1-5 6
• 2-1-6 4
• 2-1-4 5
• 2-5-4 12
(II2)
• 3-4-2 4
• 1-3-2 6
• 1-3-3 10
• 2-1-3 12
• 3-4-3 2
• 2-5-2 4
• 1-3-1 14
• 3-6-3 17
(III) : Répétition
R4 (58)
R5 (68)
R1 (63)
R6 (70)
R2 (77)
R3 (70)
ou : réussite vers réussite
: procédure erronée vers réussite
:procédure erronée vers procédure erronée

Nous exposons maintenant les idées essentielles qui ressortent de ces analyses ainsi que quelques hypothèses qu'elles nous ont conduits à formuler ; certaines sont liées aux questionnaires posés avec leurs spécificités et leurs imperfections, d'autres sont plus générales.

1. Famille (I) : mauvaise instanciation (changement de nom des objets)

C'est une classe peu stable ; l'erreur est rarement isolée ; quand elle l'est, il y a implication vers les réussites d'où notre thèse : il s'agit d'une erreur de type sémiotique, sans grande gravité.

Exception : 1-3-6 (Q_1)

L'élève écrit la bonne relation (parallélisme) de Q_1 mais entre les objets nommés dans l'hypothèse de la question suivante (Q_2) ; il s'est établi cette règle d'action ; ainsi les réponses incohérentes peuvent s'expliquer par l'invention par l'élève de règles parfois très complexes, non dénuées de bon sens mais indépendantes de la rationalité mathématique ; ici il semble que l'élève ait compris l'algorithme de la démonstration déductive à plusieurs pas, où un fait peut changer de statut d'un pas à un autre, mais non pas le sens de l'inférence : hypothèse - théorème - conclusion. A l'appui de cette interprétation, remarquons que cette procédure 1-3-6 implique la procédure 3-4-2 (Q_2) dans laquelle la conclusion donnée est strictement l'hypothèse de la question suivante.

L'application de l'algorithme semble donc l'emporter sur le traitement de l'inférence dans un problème à plusieurs pas enchaînés.

2. Famille (III) : répétition de l'hypothèse en conclusion

Dans le cas où l'erreur est isolée :

- les effectifs sont importants,
- la liaison, consistante, prouve la stabilité de ces comportements chez certains élèves,
- l'implication est très forte vers l'erreur 3-6-10 où il y a encore écho de l'hypothèse dans la conclusion.

A la faveur d'un autre test où l'unicité de la réponse n'avait pas été clairement demandée dans les consignes, nous avons observé que bon nombre d'élèves indiquaient au moins deux conclusions dont l'hypothèse de l'inférence, juxtaposant ainsi des assertions : ces élèves n'ont pas compris le sens de la forme ternaire d'un pas de démonstration. Ils considèrent la déduction comme une rhétorique du discours où la redondance est souvent pratiquée car elle assure le principe du maximum d'informations. Il s'agit d'une erreur d'ordre sémantique, stable.

3. Famille (II) : confusion entre parallélisme et égalité

A priori, la confusion peut se faire dans les deux sens («parallèle» remplacé par «égal» ou inversement) ; les deux erreurs doivent être interprétées différemment : en fait, les implications ne se font que de l'une des substitutions vers l'autre.

Nous avons pu vérifier que :

- l'erreur de remplacement de // par = n'est pas liée aux questionnaires ;
- contrairement à ce qu'on peut penser, le mot «parallélogramme» dans les hypothèses ou dans le théorème n'incite pas à écrire // plutôt que = ; le mot «milieu» n'incite pas à écrire = plutôt que // ;

- l'équivalence entre parallélogramme et quadrilatère dont les diagonales ont même milieu est très prégnante : c'est le concept «scientifique» du mot parallélogramme, celui qui est construit sous l'effet de l'apprentissage scolaire ; plus généralement, nous avons noté que les élèves articulent souvent les faits à partir de proximités des champs fréquemment et récemment conjoints ;

- pour les élèves, les droites (D), (Δ),..., ont une mobilité potentielle (au contraire, si 2, 3 points sont donnés sur une droite : droite (AB) ..., celle-ci est figée par ces points !). Par un mouvement de translation, deux droites parallèles sont superposables donc égalisables ; elles sont équivalentes en direction : dans le remplacement de droites // par droites =, il semble donc que l'égalité soit connotée d'équivalence, modulo un déplacement. L'erreur n'est pas vraiment de nature logique ; elle exprime un glissement de sens ou encore une confusion conceptuelle sans doute peu ancrée et peu persistante ;

- le remplacement de «droites parallèles» par «longueurs égales», traduisant à la fois une confusion de signes et une confusion entre un ensemble de points et un nombre, signifie certainement une erreur plus profonde ; mais moins cependant que l'erreur inverse, grave et stable ;

- deux segments distincts parallèles et isométriques sont fréquemment déclarés égaux : ils sont en effet «équivalents» en direction et en longueur (le concept de vecteur est naturellement sous-jacent à cette représentation). Deux segments isométriques mais non parallèles sont souvent aussi considérés comme égaux ; ce n'est pas le cas de 2 segments parallèles et de longueurs différentes. Au segment est donc fortement attaché un nombre : sa longueur, et l'élève écrit une égalité numérique (c'est bien ainsi qu'est perçue d'abord une égalité chez le jeune enfant).

Les analyses hiérarchiques des similarités montrent clairement la dichotomie entre les erreurs de type sémiotique et les erreurs de type sémantique, dans les questionnaires étudiés. Les analyses implicatives, elles, marquent nettement d'une part la séparation entre les confusions du parallélisme et de l'égalité (elles ne relèvent pas du même obstacle), d'autre part une différence entre les trois confusions d'objets suivantes : confusion droite et segment, confusion droite et longueur, confusion segment et longueur. Les deux hiérarchies, par leur construction même, nous informent de façon différente sans se contredire ou se complètent sans se répéter.

IV. Conclusion

Finalement, ces analyses montrent l'existence de grandes classes de procédures d'erreurs et de relations dissymétriques entre elles dans des situations de démonstration à un pas en géométrie chez de jeunes élèves. Elles confirment la stabilité de certaines conceptions, la force de certains comportements erronés et la régularité de leur apparition.

Plus précisément, elles mettent en évidence :

- **deux formes primitives de l'inférence :** la répétition, d'une part, et le changement des noms des objets avec changement ou non de la relation attendue, d'autre part ;

- **une forme plus évoluée** où les relations attendues sont échangées avec des relations voisines, sans symétrie entre ces échanges ;

- **une forme presque achevée** où la réponse diffère de l'attendu par la seule écriture (confusion entre signifiants seuls) ;
- **une forme achevée** conduisant à la réussite.

Nous espérons, dans ces quelques pages, avoir montré, en outre, la place importante que vient prendre l'implication statistique dans la panoplie des méthodes d'analyses de données, ainsi que l'opportunité et l'efficacité en didactique de ce nouvel outil. Mais son intérêt ne se limite pas, loin de là, à ce domaine. Il ouvre en effet un champ très large d'applications dans des domaines variés (évaluation par exemple) où les relations à étudier ne sont pas symétriques.

Références

AG ALMOULOUD S., 1992, L'ordinateur : outil d'aide à l'apprentissage de la démonstration et de traitement d'analyse de données didactiques, *Thèse de l'Université de Rennes I.*

GRAS R., 1979, Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques, *Thèse d'Etat, Université de Rennes I.*

GRAS R.et LARHER A, 1992, L'implication statistique, une nouvelle méthode d'analyse de données, *Mathématique, Informatique et Sciences Humaines, n° 120.*

LARHER A., 1991, Implication statistique et applications à l'analyse de démarches de preuve mathématique, *Thèse de l'Université de Rennes I.*

LERMAN I.C, GRAS R., ROSTAM H., 1981, Elaboration et évaluation d'un indice d'implication pour des données binaires, I et II, *Mathématiques et Sciences Humaines n° 74, pp. 5-35 et n° 75, pp. 5-47.*

LERMAN I.C., 1981, Classification et analyse ordinale des données, Dunod.