

Boolean experimental designs

J. Goupy

École Supérieure de Physique et Chimie Industrielles de Paris, REcherche, CONseil et FORmation, 24, avenue Perrichont, 75016 Paris. France

Experimental designs are generally interpreted using a polynomial mathematical model. But this mathematical model is not always appropriate and may sometimes not describe the phenomenon studied. Boolean experimental designs can be conceived if the factors and the response can be treated as boolean variables. The results provided are then interpreted using Boolean algebra. We have treated a real example, the settings of an instrument for analytical chemistry using both classical and boolean interpretations. The classical treatment give surprising results, with one strong interaction between two non-influent factors and interactions of order 3 and 4. The boolean interpretation gives comprehensive results and provides simple rules for the instrument settings.

Boolean modelling for the responses of an experimental design opens a new and complementary approach to the classical method that uses generally mathematical polynoms. In some cases it can provide a better interpretation of the phenomenon than the ordinary methology.

Introduction

The three types of factors involved in most experimental designs are continuous, nominal or ordinal. The response may also be continuous, nominal or ordinal. The most common mathematical model used to explain the results is the polynomial model [3,5]. This model is sufficient in most cases and it gives excellent results. However, it can sometimes lead to an erroneous or superficial interpretation. We encountered such a case in our study of an analytical chemistry method, Counter Current Chromatography (CCC). The results of an experimental design were difficult to interpret with the classical model. The main factors had no effect and there were several important interactions of orders three and four [6]. The experimental conditions suggested that Boolean variables and Boolean responses could be used. We therefore conceived a new kind of experimental design, the Boolean experimental design which proved to be efficient and gave a clear, satisfactory interpretation of the results.

Boolean variable and Boolean function

A Boolean variable can take two states: a and \bar{a} [4] such that

$$a \cup \overline{a} = 1$$
 and $a \cap \overline{a} = 0$,

where \cup is the mathematical symbol for the operation of reunion and \cap that for the operation of intersection [4].

A Boolean response, *y*, can be expressed as a function of several Boolean variables:

$$y = \varphi(a,b,c,...).$$

The function ϕ takes one of the two values: 1 or 0 depending of the variable values. The aim of a Boolean experimental design is to find out the relationship between the Boolean variables and the Boolean response.

Experimental design with two Boolean factors

Assume that a Boolean response depends on two Boolean variables. It is a classical result of Boolean algebra that two Boolean variables give $2^{2^2} = 16$ different Boolean functions [4]. There are thus 16 possible models for the response. The problem is how to identify the true model.

The low level of a factor can be attributed to the Boolean state \overline{a} and the high level to the state a. Hence, the classical geometrical representation of two level factor designs can be used. The four trials of the design are illustrated by the four corners of a square. But we must define the relationship between the two states of each variable in the Boolean design. We can use either of the two operations, intersection or reunion. The principle of the calculation is the same for the two operations. If we choose the intersection, the response at each corner of the square is given by one of these four relationships:

$$\bar{a} \cap \bar{b}$$
 $a \cap \bar{b}$ $\bar{a} \cap b$ $a \cap b$.

These are the minterms of the two Boolean variables. The results of each trial give the value for each minterm which can be 1 or 0.

In Boolean algebra, the canonical expression of a Boolean function can be written as the reunion of all the minterms

Table I. Boolean experimental matrix.

Trial #	Factor A	Factor B	Minterm	Boolean response
1	ā	\overline{b}	$\bar{a} \cap \bar{b}$	1 or 0
2	а	\overline{b}	$a \cap \overline{b}$	1 or 0
3	ā	b	$\overline{a} \cap b$	1 or 0
4	а	b	$a \cap b$	1 or 0

(disjunctive canonical form [4]). The most general model of the response is then:

$$\varphi = (\overline{a} \cap \overline{b}) \cup (a \cap \overline{b}) \cup (\overline{a} \cap b) \cup (a \cap b).$$

The experimental design gives the value of each minterm. Some minterms are zero and disappear from the function, which finally includes only the minterms equal to 1. The ultimate expression of the model can be obtained by simplifying the disjunctive canonical form. There are several methods of simplification such as those of Veitch [4] or of Harvard [4] or of Quine [9,10].

If we choose the reunion instead of the intersection to model the response at the corners of the square, we have four new relationships between the levels of A and B:

$$\bar{a} \cup \bar{b}$$
 $a \cup \bar{b}$ $\bar{a} \cup b$ $a \cup b$.

These are the maxterms of the two variables. A Boolean function can be written as the intersections of all the maxterms, and this gives the conjunctive canonical form of the response:

$$\varphi = (\overline{a} \cup \overline{b}) \cap (a \cup \overline{b}) \cap (\overline{a} \cup b) \cap (a \cup b).$$

As there is a relationship between the minterms and the maxterms, any of them can be used to treat the experimental results. A minterm, m, is the complement (negation) of a maxterm, M:

$$\overline{m}_i = M_{2^n-1-i}$$
.

We have used the minterms in the following example.

Graphical representation of a two Boolean factor experimental design

The following conventions were adopted:

The low level of A(-1) is assigned to \bar{a} .

The low level of B(-1) is assigned to \overline{b} .

The high level of A(+1) is assigned to a.

The high level of B(+1) is assigned to b.

Then the minterm $\overline{a} \cap \overline{b}$ is represented by the point [-1, -1]; the minterm $a \cap \overline{b}$ by [+1, -1]; $\overline{a} \cap b$ by [-1, +1] and $a \cap b$ by [+1, +1] (Fig. 1).

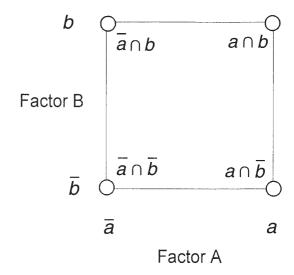


Figure 1. Illustration of a 2 factor boolean experimental design.

A similar representation was used for the maxterms. In this case, each corner of the square represents one of the expressions:

$$(a \text{ or } \overline{a}) \cup (b \text{ or } \overline{b}).$$

General Boolean designs

If we assume that a Boolean response depends on n Boolean variables, there are 2^{2^n} different Boolean functions and 2^n minterms (and the same number of maxterms). The experimental design provides the values of all the minterms that are in the disjunctive canonical form of the Boolean function that represents the response. This Boolean function generally contains many terms and must be simplified (see the following example).

The experimental design for n Boolean factors can be illustrated by an n-dimensional square. Each corner is a minterm (or a maxterm), whose value is given by the experiment. The modeling of the response is obtained from the reunion of all the minterms equal to 1. For four variables (or factors), the response at each corner is given by one these minterms:

$$(a \text{ or } \overline{a}) \cap (b \text{ or } \overline{b}) \cap (c \text{ or } \overline{c}) \cap (d \text{ or } \overline{d}).$$

There are $2^{16} = 65536$ possible functions. The minterms equal to 1 are used to write the disjunctive canonical form of the response and the diagram of Veitch is used to simplify this disjunctive canonical form and to obtain the final model.

An *n*-factor experimental design can be illustrated by an *n*-dimensional square. Figure 2 shows the graphical representation of a 4-factor experimental design.

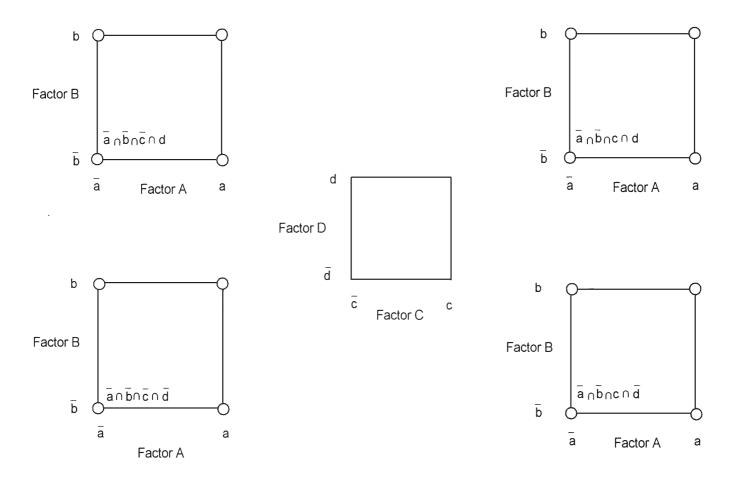


Figure 2. Illustration of a 4 factor boolean experimental design.

An example

The example is the separation method based on the partition of solutes between two immiscible liquid phases as they interact in a narrow-bore tube under centrifugal force. This method was invented by Ito [7] and is called Counter Current Chromatography (CCC) or Centrifugal Liquid-Liquid Chromatography (CLLC). Many variables or factors are involved in settings up the apparatus and this is why an experimental design was used.

Instrument description

The general principle of this apparatus is shown in figure 3 [6]. Two cylinders rotate around a vertical axis X1. The rotation can be clockwise or anticlockwise. Each cylinder also rotates on itself around the axis X2 (or X'2), which is orthogonal to X1. The two cylinders rotate at the same speed. The distance from X1 to X2 and X'2 can be adjusted by the experimenter to L or 1.5 L.

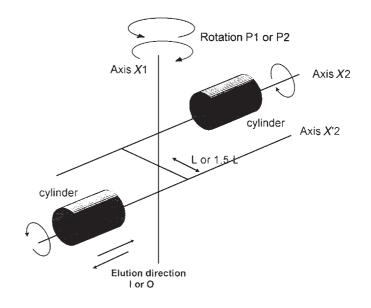


Figure 3. Simplified description of the instrument of CCC.

The chromatographic column is a narrow tube wound on the two cylinders. The mechanical arrangement is such that the tube can rotate with the cylinders without any twisting. The tube is left-handed or right-handed on the cylinders according to the instrument. When the cylinders rotate around their own axes, the ends of the column may be labeled "head" or "tail", depending on the direction of winding, (left-handed or right-handed). The choice of the "Head" or "Tail" mode is also called the "Elution mode".

The stationary liquid phase is first placed in the chromatographic column so as to occupy all the volume. Then, a mobile liquid phase (which is immiscible with the stationary phase) is passed through the column and leaves the apparatus after being in close contact with the stationary phase. The aim of the experimenter is to keep balance between the two phases in the column, so that the ratio is between 20/80 and 80/20.

The mobile phase can be heavier or lighter than the stationary phase.

The mobile phase can be introduced to flow towards the central axis X1 (Inward), or in the opposite (Outward).

The experiment

Response

The response is the percentage of stationary phase retained in the chromatographic column, as a percent of the total column volume. This percentage is called the retention of the stationary phase.

Factors

Ten factors were studied but we have select the following five:

Factor A: Density of the mobile phase. Light or Heavy.

Factor B: Elution direction. Inward (I) or Outward (O).

Factor C: Direction of rotation. Clockwise (P2) or Anticlockwise (P1)

Factor D: Cylinder position. L or 1.5 L.

Factor E: Elution mode. Tail-to-head (T) or Head-to-tail (H)

Other factors include the radii of the cylinders, the rotational speed of the cylinders, the internal diameter of the coiled tubing, the temperature and the nature of the solvent system.

Experimental design

We used a full factorial experimental design 2⁵ and considered it to be a Boolean design by choosing the intersection as operation between the Boolean states. The levels of each trial could then be transformed into a minterm. If we have chosen the reunion, each trial could have been transformed into a maxterm.

Table II. Experimental matrix and results of the example.

	=xp0:					
Trial #	Α	В	С	D	Ε	Retention
1	_	_	_	_	_	38.1
2	+	_	_	_	_	11.2
3	_	+	_	_	_	13.6
4	+	+	_	_	-	34.7
5	_	_	+	_	_	36.3
6	+	_	+	_	_	7.2
7	_	+	+	_	_	11.2
8	+	+	+	_	_	37.6
9	_	_	_	+	-	30.7
10	+	_	_	+	_	6.4
11	_	+	_	+	_	7.4
12	+	+	_	+	_	37.6
13	_	_	+	+	_	43.4
14	+	_	+	+	_	4.8
15	_	+	+	+	_	8.0
16	+	+	+	+	_	26.7
17	_	_	_	_	+	36.0
18	+	_	_	_	+	10.4
19	_	+	_	_	+	16.8
20	+	+	_	_	+	41.6
21	_	_	+	_	+	38.9
22	+	_	+	_	+	10.4
23	_	+	+	_	+	9.6
24	+	+	+	_	+	37.6
25	_	_	_	+	+	32.8
26	+	_	_	+	+	35.2
27	_	+	_	+	+	27.7
28	+	+	_	+	+	42.1
29	_	_	+	+	+	40.8
30	+	_	+	+	+	41.1
31	_	+	+	+	+	33.3
32	+	+	+	+	+	52.0
_	light	I	P2	L	Т	
+	heavy	0	P1	1,5L	Н	

The model of the response is just one of the $2^{32} = 4294967296$ possible Boolean functions.

Results

The results are shown in table II.

Classical interpretation

The classical interpretation assumes that the response can be described by a polynomial model, the coefficients of which are determined with the least square hypothesis

$$\hat{a} = (X^t X)^{-1} X^t y.$$

Table III. Classical effects and interactions.

Effects and interactions	Numerical Values	Effects and interactions	Numerical Values
		ABC	0.71
Mean	26.91	ABD	-2.13
		ABE	-2.44
Α	0.37	ACD	-0.84
В	0.43	ACE	0.79
С	0.52	ADE	1.33
D	2.46		
E	4.73	BCD	-0.37
		BCE	0.08
AB	11.02	BDE	0.16
AC	-0.63	CDE	0.99
AD	0.99		
AE	1.78	ABCD	-0.14
BC	-0.86	ABCE	0.07
BD	-0.46	ABDE	-2.65
BE	0.51	ACDE	0.96
CD	1.37	BCDE	1.36
CE	0.80		
DE	4.02	ABCDE	0.16

These coefficients are the effects of the factors and their interactions in a two level full factorial design. It is generally assumed and generally verified that the interactions of order three and higher are insignificant. But this is not so in this experimental design. Table III indicates the value of the effects and of the interactions. Some results are unusual and even surprising:

• Factors A and B have no effect and there is a very strong interaction between them.

- There are two significant interactions of order 3: ABD, ABE.
- There is one significant interaction of order 4: ABDE.

Those results allowed us to set up the apparatus and to choose satisfactory running conditions but we could not give a simple, clear interpretation of the behavior of the instrument. This is why we tried a Boolean interpretation of these experimental results.

Boolean interpretation

Each of the five factors was considered to be a Boolean variable (Tab. IV). We chose the reunion as operation between the states of these Boolean variables. Then every trial will give the value (1 or 0) of the corresponding minterm.

Table IV. Definition of boolean level of the factors.

Factor	Level -	-	Level +	
Density of the mobile				
phase (A)	light	ā	heavy a	
Elution direction (B)	inward	\overline{b}	outward b	
Direction of rotation (C)	clockwise	\overline{c}	anti-clockwise c	
Coil position (D)	L	d	1.5L <i>d</i>	
Elution mode (E)	Tail-to-head	\overline{e}	Head-to-tail e	

The response can also be considered to be a Boolean function. This function takes the value 0 when the column is empty and +1 when it is full. The response is transformed to estimate when the column can be considered to be empty or full.

Response transformation

All the responses of the experimental design (Tab. II) were recorded, from the lowest to the highest. The first rank was

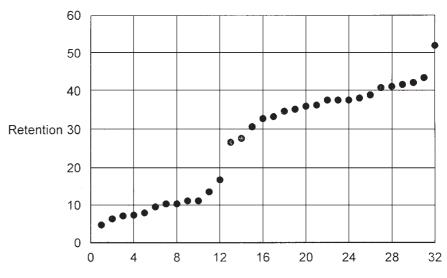


Figure 4. Retention in function of the rank of increasing values of the response.

allocated to the lowest response (4.8), the second to the next (6.4), and so on. The last rank (32nd) contains the highest response (52) (Fig. 4).

This figure shows that there is a break between the 12th and 13th responses. The retention jumps from 16.8 to 26.7. The exchanges between the two phases are excellent when there is more than 20 % of stationary phase in the column. Hence, the column could be considered to be empty when there is less than 20 % stationary phase in the column, and full when there is more than 20 % stationary phase.

The response is described by a Boolean function that has the value:

- 0 when there is less than 20 % stationary phase in the column.
- +1 when there is more than 20 % stationary phase in the column.

The results of the initial design (Tab. II) are transformed into Boolean responses according to these statements (Tab. V). This table shows the minterms of the disjunctive canonical form of the response.

Table V. Boolean experimental design and results of the example.

Trial #	А	В	С	D	E	Minterm	Boolean response
1	ā	b	\overline{c}	ā	ē	$\overline{a} \cap \overline{b} \cap \overline{c} \cap \overline{d} \cap \overline{e}$	1
2	а	\overline{b}	\overline{c}	\overline{d}	ē		0
3	ā	b	\overline{c}	\overline{d}	ē		0
4	а	b	\overline{c}	\overline{d}	ē	$a\cap b\cap \overline{c}\cap \overline{d}\cap \overline{e}$	1
5	ā	\overline{b}	С	\overline{d}	ē	$\overline{a} \cap \overline{b} \cap c \cap \overline{d} \cap \overline{e}$	1
6	а	\overline{b}	С	d	ē		0
7	ā	b	С	\overline{d}	ē		0
8	а	b	С	\overline{d}	ē	$a\cap b\cap c\cap \overline{d}\cap \overline{e}$	1
9	ā	$\frac{\overline{b}}{b}$	\overline{c}	d	ē	$ar{a}\capar{b}\capar{c}\cap d\capar{e}$	1
10	а	\overline{b}	\overline{c}	d	ē		0
11	ā	b	\overline{c}	d	ē		0
12	а	b	\overline{c}	d	ē	$a \cap b \cap \overline{c} \cap d \cap \overline{e}$	1
13	ā	\overline{b}	С	d	ē	$\overline{a} \cap \overline{b} \cap c \cap d \cap \overline{e}$	1
14	а	\overline{b}	С	d	ē		0
15	ā	b	С	d	ē		0
16	а	b	С	d	ē	$a\cap b\cap c\cap d\cap \overline{e}$	1
17	ā	\overline{b}	\overline{c}	\overline{d}	е	$ar{a}\capar{b}\capar{c}\capar{d}\cap e$	1
18	а	\overline{b}	\overline{c}	a a a a a a a	е		0
19	ā	b	\overline{c}	\overline{d}	е		0
20	а	b	\overline{c}	\overline{d}	е	$a \cap b \cap \overline{c} \cap \overline{d} \cap e$	1
21	ā	$\frac{\overline{b}}{b}$	С	\overline{d}	е	$\bar{a} \cap \bar{b} \cap c \cap \bar{d} \cap e$	1
22	а	\overline{b}	С	\overline{d}	е		0
23	ā	b	С	\overline{d}	е		0
24	а	b	С	\overline{d}	е	$a \cap b \cap c \cap \overline{d} \cap e$	1
25	ā	\overline{b}	\overline{c}	d	е	$\overline{a} \cap \overline{b} \cap \overline{c} \cap d \cap e$	1
26	а	\overline{b}	\overline{c}	d	е	$a\cap \overline{b}\cap \overline{c}\cap d\cap e$	1
27	ā	b	\overline{c}	d	е	$\bar{a} \cap b \cap \bar{c} \cap d \cap e$	1
28	а	b	\overline{c}	d	е	$a \cap b \cap \overline{c} \cap d \cap e$	1
29	ā	\overline{b}	С	d	е	$\bar{a}\cap\bar{b}\cap c\cap d\cap e$	1
30	а	\overline{b}	С	d	е	$a\cap \overline{b}\cap c\cap d\cap e$	1
31	ā	b	С	d	е	$\bar{a} \cap b \cap c \cap d \cap e$	1
32	а	b	С	d	е	$a \cap b \cap c \cap d \cap e$	1
	P 1 4		D2				
factor	light	I	P2	L	T		
factor	heavy	0	P1	1,5L	Н		

factor	light	I	P2	L	Т
factor	heavy	0	P1	1,5L	Н

If the experimenter wants, for example, a higher retention, he can choose a different cut off. The only difference is that the Boolean function will not be the same.

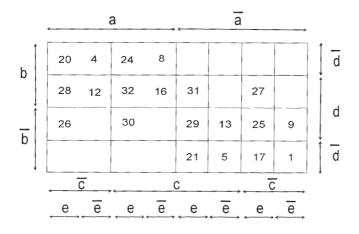


Figure 5. Veitch diagram for five factors.

The Boolean disjunctive canonical form

Table V shows the minterms corresponding to each trial. The disjunctive canonical form of the response is the reunion of the 20 minterms whose values are 1 in the response column of the Boolean experimental design.

$$\varphi = (\overline{a} \cap \overline{b} \cap \overline{c} \cap \overline{d} \cap \overline{e}) \cup (a \cap b \cap \overline{c} \cap \overline{d} \cap \overline{e}) \cup ... \cup (a \cap b \cap c \cap d \cap e).$$

This is a complicated function that must be simplified.

Simplification of the Boolean function

A five-variable Veitch table was used [4] (Fig. 5), in which each cell is a minterm corresponding to one trial of the experiment. There are $2^5 = 32$ minterms.

The number of the trial is inserted in the corresponding cell when the value of a minterm is 1. The numbered cells allowed us to establish the simplified Boolean function of the response. The eight cells 4, 8, 12, 16, 20, 24, 28 and 32 are due to the intersection of a and b: $a \cap b$. Cells 25, 26, 27, 28, 29, 30, 31,and 32 are due to the intersection of d and e: $d \cap e$. Cells 1, 5, 9, 13, 17, 21, 25 and 29 are due

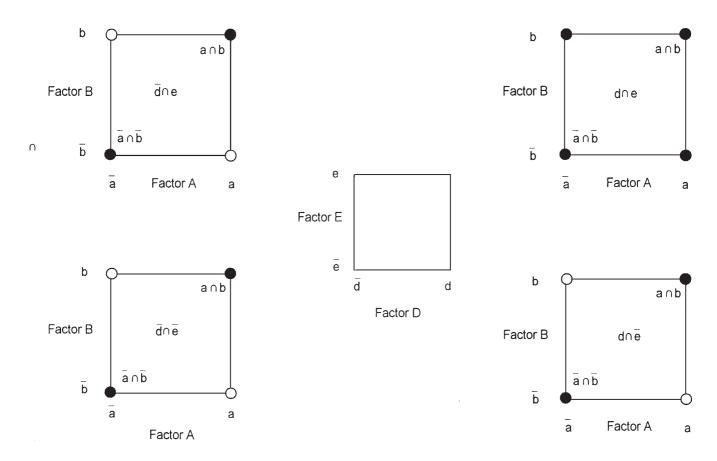


Figure 6. Black dots show all satisfactory settings for the CCC instrument: the stationary phase is immobilisied in the chromatographic column.

to the intersection of \overline{a} and \overline{b} : $\overline{a} \cap \overline{b}$. The resulting simplified function is

$$\varphi = (a \cap b) \cup (\overline{a} \cap \overline{b}) \cup (d \cap e).$$

This function is readily interpreted:

- Factor C has no influence on the response.
- The column will be full if only one of the intersections is equal to 1. This means that the setting of the instrument could be one of the following:
 - D and E must be set together at high level, regardless of the levels of the other factors.

The "L" cylinder position must be used with "Head to tail" (H) elution mode;

01

 A and B set together at high level, regardless of the levels of the other factors.

A "high density" mobile phase needs the "Outward" elution direction;

or

A and B set together at low level, regardless of the levels of the other factors.

The "low density" mobile phase needs the "Inward" elution direction.

Figure 6 shows the combinations of factor levels that give a full column (retention above 20 %) and the combinations that give an empty column.

Conclusion

The classical polynomial model generally used to interpret the results of an experimental design is not suitable for all problems. We have used Boolean functions to interpret the results of an experimental design. Sometimes these functions can be more appropriate than the classical polynomial functions. Boolean functions need to introduce a relationship between the factor levels. This relationship can be the "reunion" or the "intersection". Then, the mathematical model of the response at each experimental point is given by a minterm (or a maxterm). The experimental results give values of 1 or 0 for each minterm. The disjunctive canonical form is obtained by the reunion of all the mintern of value 1. The disjunctive canonical form is simplified to a model that can be easier to understand and more practical. The example of CCC instrument settings shows how Boolean experimental designs can be used.

References

- 1. Boole, G. *The mathematical analysis of logic*; Cambridge, 1847, republished by Blackwell, 1948.
- 2. Boole, G. An investigation of the law of thought; London, 1854, republished by Dower Publications.
- 3. Box, G.E.P.; Hunter, W.G.; Hunter, J.S. Statistics for Experimenters; New York: John Wiley and Sons, 1971, 453 pages.
- 4. Denis-Papin, M.; Faure, R.; Kaufmann, A. Cours de calcul booléien; Paris: Éditions Albin Michel, 1970, 495 pages.
- 5. Goupy, J. Methods for Experimental Design. Principles and Applications for Physicists and Chemists; Amsterdam: Elsevier, 1993, 449 pages.
- Goupy, J.; Menet, J.-M.; Shinomyia, K.; Ito, Y. in *Modern Countercurrent Chromatography*; Conway, W.D.; Petroski, R.J. (Eds) ACS Symposium Series 593, Washington DC.: American Chemical Society, 1995, Chapter 4, pp 47-61.
- 7. Ito, Y.; Mandava, N.B. Countercurrent chromatography. Theory and practice; New York: Marcel Dekker, 1988.
- Ito, Y.; Zhang, T-Y. Journal of Chromatography 1988, 449, 135-151.
- 9. Quine, W. V. O. The American Math. Monthly 1952, October.
- 10. Quine W. V. O. *The American Math. Monthly* **1955**, November.
- 11. Shinomyia K., Menet J.-M.; Ito, Y. Journal of Chromatography 1993, 44, 215-229.
- 12. Goupy, J. *Plans d'expériences pour surface de réponse*; Paris: Dunod, 1999, 409 pages.
- Menet, J.-M.; Thiébaut, D. Countercurrent chromatography; Chromatographic Science Series Volume 82, New York: Marcel Dekker Inc., 1999, 279 pages.