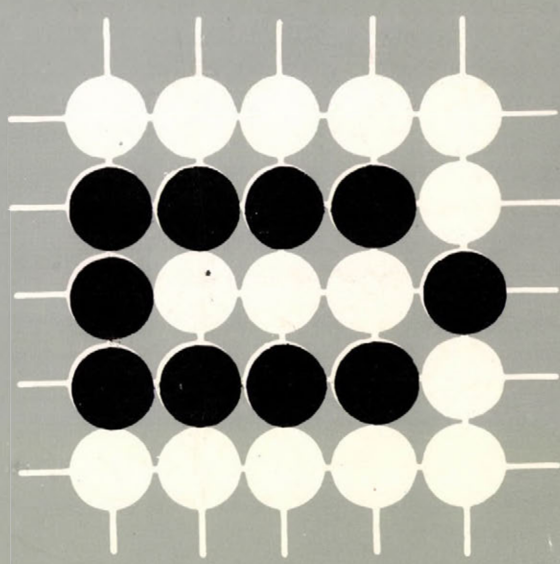
közlemények

39/1988

MTA Számítástechnikai és Automatizálási Kutató Intézet Budapest



Magyar Tudományos Akadémia
Számitástechnikai és Automatizálási Kutató Intézete
Computer and Automation Institute, Hungarian Academy of Sciences

K Ö Z L E M É N Y E K

T R A N S A C T I O N S

39/1988

Felelős Kiadó:

KEVICZKY LÁSZLÓ

ISSN 0133-7459
ISBN 963 311 253 2

CONTENTS

C O N T E N T S

MTA SZTAKI Közlemények/39 p. 5-26.

DEFINING A CONNECTION FUNCTION AS A BASE FOR A USER INTERFACE TO A RELATIONAL DATABASE

Zh. S. ANGELOV

"Interprograma"
P.o.b. 795
1000 Sofia, Bulgaria

Introduction

The wide proliferation of computers in information servicing puts forward the problem for the development of user interfaces (UI), i.e. the methods for the mapping of queries, formulated in terms of the problem area into queries, formulated in terms of the system storing and maintaining data. Such a problem arises in the development of information systems based on DBMS. The DBMS's used nowadays support data models, which offer simple structures for the modelling of the real world and are oriented to easier physical representation. In order to avoid redundancy the designer cannot capture in the schema all known semantic relationships from the problem area when modelling with the help of these structures. The end-user, who is not a specialist in data processing but is a specialist in the problem area and knows these relationships will probably use them formulating his queries. A problem area with a given semantics which is known to any end-user can often be modelled in different ways, for example depending on the frequency of queries to DBMS. Therefore we need a way for mapping the user view about the part of the modelled area into operations over the structures of the stored data. One possibility is to define a connection function. This function is a mapping of the database state into relations over sets of the problem area terms. We can say that the connection function has as its values the usual user views. But user views in DBMS must be defined in advance for any set of attributes, while the connection function, once defined, gives us a way to compute these views for any subset of problem area terms.

This paper discusses the assumptions under which a connection function can be defined. Its properties and the way for its definition are also discussed. An algorithm, based on the connection function for the development of UI, is proposed.

Assumptions

The proposed approch for building an end-user interface assumes that the attributes carry the whole semantics of the problem area and that the relational schemes are constructed taking into consideration information processes and are not uniquely determined by the diven problem area. Hence, from user point of view the data describing the problem area are stored in a single relation over the set of all attributes. This is the so called universal relation (UR).

In order to ensure the adequacy of the idea of UR existence and the possibility of its automatical maintainance the database schema should possess some properties. For this reason

In the works adopting UR approch some assumptions about the database scheme are made [Ang85, Ang86b, FMU82, MRW83, MUV84, Men84, Osb79, Roz83, Sag83, Ull83].

The first basic assumption, which all authors accept, is:

Assumption 1. 'Universal Relation Scheme Assumption' (URSA). Any attribute in U corresponds to the same class of entities wherever it appears.

Most of the authors understand under this assumption the fact that any attribute plays only one role. For example, NUMBER cannot refer to the number of children and to the department number of an employee.

The next assumption discusses the connection among the given set of attributes $X \subset U$. In order to be able to bind together relations automatically, i.e. to bind together different attributes there should be set up a basic semantic relationship in the scheme. The user should have in mind the same relationship when thinking about the attributes of the given problem area.

Assumption 2. 'Relationship Uniqueness Assumption' (RUA). Let $X \subset U$. There exists only one basic semantical relationship among the attributes X. The user means this unique relationship (denoted [X]) when talking about the attributes X as a whole.

The relationship between the attributes in X={TEACHER,STUDENT} is an example of the uniqueness of the basic semantic relationship. Speaking about a teacher and a student together, first of all we have in mind that "The teacher teaches the student".

Let us consider an example taken from [MU83] and fairly often used in cases when the adequacy of one or another formalism, describing semantic relationship among attributes is discussed [Ang86a,D'AMS83, MRW83, MU83, Roz83].

Example 1. Let us consider a banking database. The attributes are BNK (bank), ACC (account), L (loan), C (customer), AMT (loan amount), BAL (account balance) and ADR (customer address), i.e. U={BNK,ACC,L,C,AMT,BAL,ADR}. Let the database scheme be $D=\{R_1,R_2,\ldots,R_7\}$, where the relational schemes are respectively $R_1=\{C,L\}$, $R_2=\{C,ACC\}$, $R_3=\{C,ADR\}$, $R_4=\{AMT,L\}$, $R_5=\{L,BNK\}$, $R_6=\{BNK,ACC\}$ и $R_7=\{ACC,BAL\}$. The hypergraph of the database scheme is depicted on fig.1. □

If $X_1=\{C,ACC\}$, then $[X_1]$ means "<u>The customers</u> own <u>accounts</u>". If $X_2=\{C,ACC,BNK\}$, then $[X_2]$ means "<u>The customers</u> own <u>accounts</u> at <u>the banks</u>". Similarly for $X_3=\{C,L,BNK\}$, $[X_3]$ means "<u>The customers</u> have taken out <u>loans</u> from <u>the banks</u>". If we consider $X_4=\{C,BNK\}$, then under $[X_4]$ we must understand "<u>The customers</u> are served at <u>the banks</u>" or "<u>The banks</u> serve <u>the customers</u>". Finally, if $X_5=\{ACC,L\}$, then $[X_5]=\varnothing$, because there is no basic semantic relationship (The relationship is neither "the loans and accounts at one and the same bank", nor "the loans and accounts of one and the same customer").

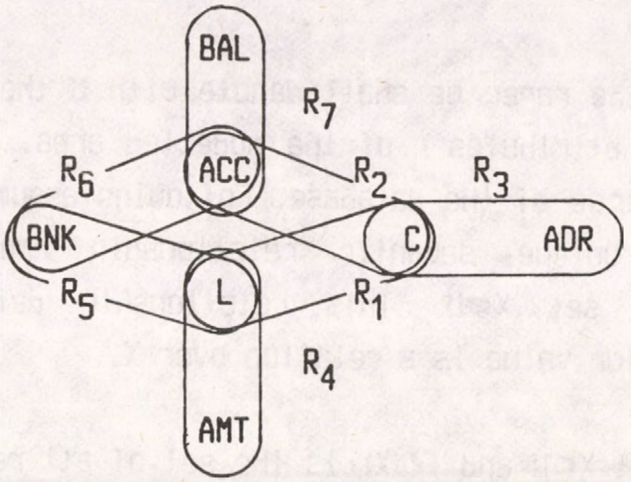


Fig. 1. The banking database scheme

Let us consider again the set X={C,BNK}. If we connect the attributes C and BNK (see fig.1.) the access path can pass through the attribute ACC (i.e. the customers own accounts at the given banks, so they are served by them) or pass through the attribute L (i.e. the customers have taken out loans from the given banks, so they are again served by them). In this case two access paths are possible but both have the same 'flavor' (the customers are served by the bahks).

The possibility to connect attributes through more then one path arises when the database scheme is cyclic as the one in example 1. In this case the database scheme must satisfy the following assumption:

Assumption 3. 'One Flavor Assumption' (OFA). All access paths used to compute the connection on X represent the same 'flavor' of the relationship among the attributes in X.

Connection function and its properties

In the rest of the paper we shall denote with U the set of all terms (called attributes) of the modelled area. This is the so called **universe** of the database. Following assumption 2 there exists an unique semantic relationship among the attributes of the set $X \subset U$. This relationship defines a function. The function value is a relation over X.

Definition 1. Let $X \subset U$ and $\mathbf{R}(X)$ is the set of all relations over X. A function $[X](d)$ is a **connection function** if it maps the database state d to the set $\mathbf{R}(X)$. (We will omit (d) and denote only with [X] the value of the connection function. The functional [.] maps a subset of U to a function from database states to relations over that subset.)

The connection function is named also simply "connection" [FMU82] and "window function" or "window" [MRW83].

The use of the connection function has some effects which lead to the requirement that the connection function should possess certain properties. Such a property is the satisfaction of the containment condition.

Definition 2. [MRW83] The connection function [.] satisfies the **containment condition** if the inclusion $X \subset Y \subset U$ implies $\pi_X [Y] \subset [X]$.

In other words the following principle must be reflected in the connection function: When one speaks in general (considering only a few attributes) one refers to more objects than when specifying more details (i.e. when one is interested in more attributes).

Let us consider the database from example 1. The containment condition implies the fact: "The customers served at the banks as a whole" are not less than "the customers who own accounts at the banks" (i.e. $\pi_{\{C,BNK\}}[\{C,L,BNK\}] \subset [\{C,BNK\}]$).

When asking for some data the end-user expects to retrieve the data which has been inserted, i.e. if the user has added a tuple t to a stored relation $r(R)$ then the value of the connection function for the set R must contain such a tuple t. In other words the connection function must provide visibility of all tuples stored in the database. This property is called "faithfulness".

Definition 3. [MRW83] The connection function is **faithful** (possesses the property **faithfulness**) if $r(R) = [R]$ holds for any relational scheme $R \subseteq U$ and any database state d.

This means that the connection function must neither hide tuples nor add any. Therefore we have to explicitly store in the database all known facts. Taking into consideration the trend to introduce deductive capabilities in DBMS, a more realistic definition of this property may be $r(R) \subset [R]$.

Approaches to the definition of the connection function

All known methods for the definition of a connection function include joins of the stored relations. Relations are the minimal objects which can be updated. Therefore the relational schemes can be called **update structures.** These are the base for the conection function definition. Following the fact that a "good" connection function must satisfy the containment condition and is a monotonously decreasing function we can construct for any definition method a special kind of structures - **retrieval structures.** These structures are sets of attributes.

Definition 4. The set of attributes $X \subset U$ is a retrieval structure if there exists a database state such that the connection function value is not the empty set ($[X] \neq \varnothing$) and the extension of X with any other attribute $A \in U-X$ does not have this property (i.e. $[X \cup \{A\}] = \varnothing$).

As the retrieval structure coincides with the union of some of the update structures we can consider the retrieval structures as sets of update structures.

The retrieval structure semantically corresponds to one of the possible aspects of the basic semantic relationship. Thus the definition of the connection function can be formulated as follows:

Definition 5. Let $V-\{R_1, R_2, \ldots, R_n\}$ is the set of the update structures and $W=\{S_1,S_2,\ldots,S_m\}$ is the set of retrieval structures, where $S_i \subset V$. Then the connection function is

$$[X] = \bigcup_{S\in W, X\subseteq attr(S)} \pi_X(\bowtie_{R\in V, R\in S} r(R)),$$

where $attr(S) = \bigcup_{R\in S} R$.

Therefore, the way of constucting the retrieval structures is significant for the connection function definition. There exist two approaches of using (i.e. constructing) the retrieval structures:

- the structures are not explicitly given and have to be constructed from the database scheme using some additional information as functional dependencies (for example, building "lossless joins", in particular "extension joins");

- the structures are explicitly given, i.e. a set of new structures is added to the database scheme (for example "maximal objects" in System/U or "objects" in PIQUE).

The nonexplicitly defined retrieval structures are based on one or more classes of data dependencies (mainly functional ones). As a consequence if the modelled area is more complex and its semantics cannot be described using these kinds of dependences, we cannot obtain an adequate connection function definition. This gives us confidence to claim that the explicit definition gives better results.

The explicit definition raises some problems too. It is not absolutely clear how to create the retrieval structures. We can use some algorithms to construct possible retrival structures, but sometimes the obtained structures may be not very good. The example in [KKFGU84] shows that little changes in the modelled area can imply essential reconstruction in the set of retrieval structures.

These difficulties show that the database scheme and integrity constraints are not sufficient for the adequate definition of retrieval structures which are used in UI. This is a consequence of the loss of some information in the process of the mapping the conceptual model of the problem area to model supported by the DBMS. All proposed methods discuss the creation of retrieval structures after defining the database scheme, i.e. after defining the update structures. The process is depicted on fig.2. Thus the way in which the database scheme

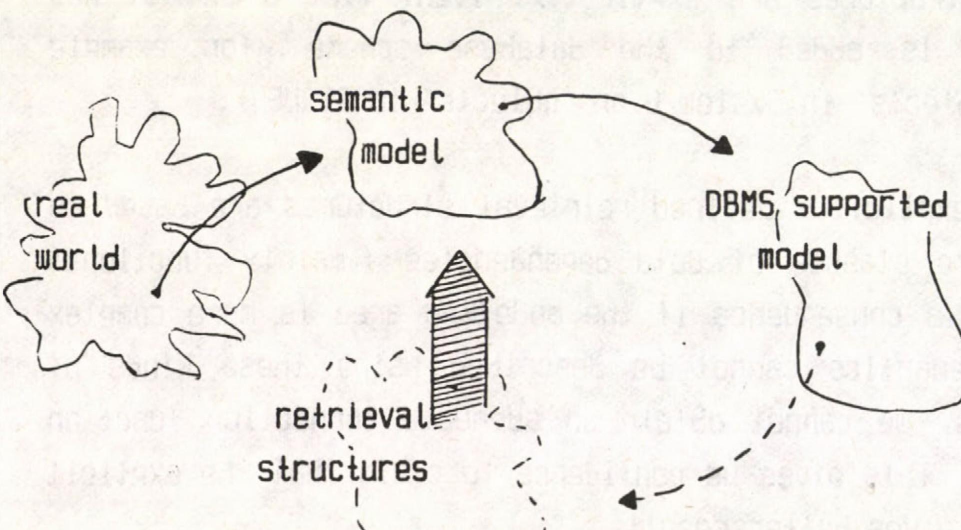


Fig.2. Mapping the models in the database scheme design process

is obtained, is lost. If we (in a semantic or a conceptual model) define first the retrieval structures and after that map them onto the update structures, then we automatically obtain a definition of a connection function. In this way the last mapping in fig.2. may be removed and we have the chance to capture more meaning. This approach is followed in [Ang86a, Ang86b]. The retrieval structures are constructed using the aggregation hierarchy.

An algorithm for an user interface to a DBMS, supporting an universal relation as an user view

The query language used by the end user has to provide the following two capabilities:

- explanation of the target attributes;
- specification of the conditions which the target data must satisfy.

If the query language is intended to free the user from the logical navigation then it obeys the following principle: The sentences of that language cannot include any structure related to data storage details. Instantiating this principle in the relational data model we obtain the assertion that the query language cannot include any relation names. Thus a query may consist only of attribute names and the way they are related. In this case the user wants to extract information from a view. This view must include all attributes mentioned in the query and it can be calculated using a connection function. In this way for any query we can juxtapose a value of the connection function for the set of attributes mentioned in that query. The tuples which contain the information of interest are among the tuples of the connection function. The former can be extracted through testing which of the latter satisfy the conditions in the query.

In almost every known system, based on the universal relation scheme, the query processing is separated into two steps [Ang86b, KKFGU84, MRS85, MRW83, MUV84]:

a) binding (i.e. user view creation). It consists of the construction of the connection function [X] for the set X of attributes mentioned in the query;

b) evaluation (i.e. target data extraction). Whatever operations must be applied to answer the query, are then applied to [X].

The following example illustrates these two staps.

Example 2. Let us consider the database from example 1 and let d be a database state. Let the connection function [X] be defined as a projection of joins of relations with relational schemes covering the set of attributes X, i.e. $[X]=\pi_X(\bowtie_{R\in M} r(R))$, where $X\subset attr(M)$ and $r(R)\in d$. If the user query is "Find all account balances of the customer named 'Angelov'", then the answer can be contructed in the afore said two steps. The first step is the binding of the attributes mentioned in the query, namely {BAL,C}. In other words, we build an expression whose value gives the value of the connection function for {BAL,C}. According to the above definition the expression is $[\{BAL,C\}]=\pi_{\{BAL,C\}}(r(R_2)\bowtie r(R_7))$. The second step should reflect the fact that the user would like to retrieve information referring to the customer named 'Angelov' only. Therefore we have to use the selection operation over [{ BAL,C }]. Finally, the answer to the user query is the value of the expression

$$\sigma_{C='Angelov'}\pi_{\{BAL,C\}}(r(R_2)\bowtie r(R_7)) \qquad \square$$

Setting up the construction of a connection as a first step allows us to test easily new ideas for the building of the connection function (as in [Ang86b]). We would like to stress that this gives also the opportunity to use different algorithms depending on the available additional information. Below we will describe a generalized algorithm for an user interface to a RDBMS (without taking into consideration the type of the additional information). This algorithm allows the user to view the data in the database as stored into one unique relation. But firstly let us name some of the sets and relations, which will be used later: MENSET will denote the set of all attributes mentioned in the query, ANSSET - the set of the target attributes and $RETSTR_i$ - the set of relational schemes included in a given retrieval structure. RETSTRSET is the set of all sets $RETSTR_i$ which cover MENSET. The relation of the user view will be denoted by **window** and the answer will be received in the relation **answer**. **window** is a relation over MENSET and **answer** is a relation over ANSSET. Let **cond** be the condition formulated in the given query. As a query language we can assume a modification of SQL according to the principle stated above, i.e. only the use of attribute names is permitted.

Algorithm 1. Construction of the answer to a query to a system, supporting an universal relation.

Input: An user query specified in "modified" SQL.

Output: The relation **answer** which is the answer to the user query or a message for the "meaninglessness" of the query if the set MENSET is unconnected.

Method:

1. Extract the names of all the attributes mentioned in the user query, thus constructing the set MENSET. Extract the names of the target attributes and construct the set ANSSET. Construct the expression **cond**.

2. Find all retrieval structures $RETSTR_i$ covering the set MENSET (i.e. $MENSET \subset attr(RETSTR_i)$).

3. If there exists no retrieval structure which covers MENSET (i.e. RETSTRSET-∅), then output the message for the "meaninglessness" of the query and stop.

4. For each i reduce the retrieval structure $RETSTR_i$. Remove any set $RETSTR_i$ which is a superset for any other set $RETSTR_j$.

5. For each i compute the join of the relations over schemes from $RETSTR_i$. Project the result over MENSET and make its union with the temporary relation **window**.

6. Remove (using selection) all tuples which do not satisfy the condition **cond**.

7. Project the relation **window** over ANSSET in order to obtain the relation **answer**.

Step 1 includes only preliminary procedures. In step 2 the connection function is formed. Here are the main differences among the various approaches. As mentioned above, different algorithms can be used depending on the available information.

Step 3 has a control function. The procedures included in step 4 aim to optimize the creation of the relation **window** (the value of the connection function). Although most of the definitions of the connection function take into consideration some criteria for optimization, this leads to local optimization only. Generaly, it is assumed that the optimization is carried out by the DBMS and not by the interface. The reduction of a retrieval structure aims to remove relations which will not change the result. The information of interest, contained in those relations is included into the remaining relation. In other words, the remaining relations contain more general information about the particular case. As an example let us consider the expression

$$[\{C,L,BNK\}] - \pi_{\{C,L,BNK\}} (r(R_1) \bowtie r(R_3) \bowtie r(R_5) \bowtie r(R_4))$$

(Note that the definition of [.], given in example 2., is used here). The relation $r(R_3)$ contains information for all bank customers. The attribute L in the set {C,L,BNK} shows that we are interested only in the customers which have taken out loans. The absence of the attribute ADR shows that we are not interested in the customers' addresses. Therefore the inclusion of the relation $r(R_3)$ in the evaluation of the expression will be a pure loss of time. The reduction can be performed in different ways. For example, in System/U minimization under weak equivalence is used. A different method is given below, which is a modification of the Graham reduction. The proposed second part of step 4 follows directly from the join property. Step 5 calculates the value of the connection function. Step 6 and 7 perform the second part of query processing.

It should be pointed out that step 4 is optional. If step 4 is omitted, the result will be the same but in this case we will lose much more resources than are needed for the execution of step 4. That is why we suggest to use Graham reduction algorithm. The original Graham algorithm [Gra80] is intended for testing hypergraph acyclicity. For the purpose of optimization we propose a modified algorithm. The idea is: Consider the attributes mentioned in the query as a goal set and the hypergraph, having as its nodes the attributes and as edges - the relational schemes from this structure. The algorithm removes edges and nodes preserving the goal set and the hypergraph connectivity.

Algorithm 2. Reduction of a set of relational schemes with respect to a goal set of attributes.

Input: A set of relational schemes $\mathbf{R}$ and a goal set of attributes $\mathbf{G} \subset attr(\mathbf{R})$.

Output: A set of relational schemes $\mathbf{R}' \subset \mathbf{R}$ which when considered as a hypergraph has the following property: for any two elements of $\mathbf{G}$, there exists a path between them.

Method:

1. Let $\mathbf{R}' - \mathbf{R}$.

2. Apply one of the following operations as many times as possible over the relational schemes in $\mathbf{R}$:

a) If A is not in $\mathbf{G}$ and is included in only one relational scheme $R_i \in \mathbf{R}$ then remove A from R_i;

b) If $R_i \in \mathbf{R}$ is a subset of $R_j \in \mathbf{R}$ then remove R_i from the sets $\mathbf{R}'$ and $\mathbf{R}$. □

Conclusion

In the paper an approach to the development of an user interface is considered. It is based on the use of the connection function. It is shown that the problems arising in development of the UI can be separated from the specific model of the problem area which is used as a database scheme. The introduction of the connection function as a base for the UI makes the latter independent from the manner of the real world modeling. This facilitates the transition towards the use of a more complex model if the one in use does not possess enough expressive power. There is a price to be paid for this freedom by the database administrator. He (or she) has to create and maintain a database scheme which satisfies the given conditions. In the paper, such conditions are formulated and discussed. Such an interface may serve as a workbench and can be an useful tool for testing new ideas in the field of the automation of query answering.

References

[Ang85] Angelov Zh.S. : Towards a Universal Relation View. In Proceedings Eighth International Seminar on Database Management Systems, Piestany, Czechoslovakia, 1985, pp. 9-17.

[Ang86a] Angelov Zh.S. : A Connection Function for End-User Interface. In Proceedings XV-th Spring conference of SMB, Sunny beach, Bulgaria, 1986, pp. 350-355. (in bulgarian)

[Ang86b] Angelov Zh.S. : Simulating a Universal Relation View

Introducing Aggregation Information. In Proceedings Ninth International Seminar on Database Management Systems, Reinhardsbrunn, GDR, 1986, pp. 189-194.

[D'AMS83] D'Atri A., Moscarini M., Spyratos N. : Answering queries in relational database. In Proceedings ACM SIGMOD Annial Meeting, San Jose, May 23-26, 1983, pp. 173-177.

[FMU82] Fagin R., Mendelzon A.O., Ullman J.D. : A Simplified Universal Relation Assumption and Its Properties. ACM Trans. Database Syst. Vol.7, No.3, 1982, pp. 343-360.

[Gra80] Graham M.H. : On the Universal Relation. In A Panache of DBMS Ideas III. (ed.) D Tsichritzis. Technical Report CSRG-111, University of Toronto, 1980, pp. 68-92.

[KKFGU84] Korth H.F., Kuper G.M., Fiegenbaum J., Gelder A., Ullman J. : System/U - A Database System based on the Universal Relation Assumption. ACM Trans. Database Syst. Vol.9, No.3, 1984, pp. 331-347.

[MRS85] Maier D., Rozenshtein D., Stein J. : Representing roles in universal scheme interfaces. IEEE Trans. on Software Engineering. Vol.SE-11, No.7, 1985, pp. 644-652.

[MRW83] Maier D., Rozenshtein D., Warren D.S. : Windows on the World. Proceedings ACM SIGMOD International Symposium on Management of data, San Jose, Calif., 1983, pp. 68-78.

[MU83] Maier D., Ullman J.D. : Maximal objects and the Semantics of Universal Relation Databases. ACM Trans. Database

Syst. Vol.8, No.1, 1983, pp.1-14.

[MUV84] Maier D., Ullman J.D., Vardi M.Y. : On the Foundations of the Universal Relation Model. ACM Trans. Database Syst. Vol.9, No.2, 1984, pp. 283-308.

[Men84] Mendelzon A.O. : Database states and their tableaux. ACM Trans. Database Syst. Vol.9, No.2, 1984, pp. 264-282.

[Osb79] Osborn S.L. : Towards a Universal Relation Interface. In Proceedings Conf. on Very Large Data Bases V, 1979, pp. 52-60.

[Roz83] Rozenshtein D. : Query and Role Playing in the Association-Object Data Model. Ph.D. dissertation, SUNY at Stony Brook, 1983.

[Sag83] Sagiv Y. : A Chacterization of Globally Consistent Databases and Their Correct Access Paths. ACM Trans. Database Syst. Vol.8, No.2, 1983, pp. 266-286.

[Ull83] Ullman J.D. : Universal Relation Interfaces for Database Systems. In Information Processing 83, R.E.Mason(ed), Elsevier Science Publishers B.V. (North-Holland), pp. 243-242.

DEFINING A CONNECTION FUNCTION AS A BASE FOR A USER INTERFACE TO A RELATIONAL DATABASE

Zh. S. Angelov

Summary

In the paper an approach to the development of a user interface is considered. It is based on the use of the connection function. The introduction of the connection function as a base for the user interface (UI) makes the latter independent from the way of the real world modelling. There is a price to be paid for this freedom by the database administrator. The administrator has to create and maintain a database scheme which satisfies the given conditions. In the paper such conditions are formulated and discussed. Such an interface may serve as a workbench and can be a useful tool for testing new ideas in the field of the automation of query answering.

KAPCSOLATFÜGGVÉNYEN ALAPULÓ FELHASZNÁLÓI INTERFACE RELÁCIOS ADATBÁZISOKHOZ

Zh. S. Angelov

Összefoglaló

A dolgozat egy kapcsolatfüggvényen alapuló, a valós világ modellezésének módjától független felhasználói interface-szel foglalkozik. Ezt a függetlenséget természetesen nem adják ingyen: az adatbázis adminisztrátorának létre kell hoznia és karban kell tartania egy bizonyos feltételeknek eleget tevő sémát. A dolgozat megadja, és részletesen megvizsgálja ezeket a feltételeket. Az interface az automatizált lekérdező rendszer területén az uj ötletek kipróbálásának hasznos eszköze lehet.

MTA SZTAKI Közlemények 39/1988. p. 27-33.

ON THE SOLIDITY OF PACKINGS OF INCONGRUENT CIRCLES I.

A. HEPPES

Computer and Automation Institute
Hungarian Academy of Science

1. Introduction.

A packing of convex discs is said to be *solid* if no finite subset of the discs can be rearranged so as to obtain a packing *not* congruent to the original one [1].

In the present paper we shall prove a general theorem that contains sufficient conditions for the solidity of circle packings in the Euclidean plane.

2. Definitions.

Let $\mathcal{R} = \{r_1, r_2, \dots, r_k\}$ be a set of positive numbers. Consider three disjoint open circles of a radius $\rho_i \in \mathcal{R}$ $(i=1,2,3)$. This triple as well as the triangle determined by the centers of the circles will be called *normal* if none of the segments connecting two centers intersects the third circle.

We say that a set of normal triangles *generates a packing* if the triangles cover the plane without gaps and without overlapping and the circle sectors of the individual normal triangles fit together to from complete circles.

A positive *weight* $w(r_i)$ will also be assigned to all circles of radius r_i $(i=1,2,\dots,k)$.

Let O_j and ρ_j denote the centers and the radii of a normal triple, respectively, and α_j the angle of the trialgle at vertex O_j $(j=1,2,3)$. The *weighted density* of the triple (in the triangle) is defined by

$$\sigma = \frac{\frac{1}{2}.\pi.\sum_{j=1}^{3}\rho_j^2.\alpha_j.\omega(\rho_j)}{A},$$

where A denotes the area of the triangle $O_1O_2O_3$.

For the sake of simplicity the term density will be used instead of weigthed density throughout this paper.

A normal triangle will be called *tight* (*spanned*) if the circles are mutually tangent (if one circle is tangent to the other two and to the opposite side) (Fig. 1).

3. Preparations.

First we show the validity of the following

LEMMA 1. Let the radii r_i and weights $\omega(r_i)$ (i=1,2,3) be given. We consider all normal triples consisting of circles the radii of which belong to the set $\mathcal{R} = \{r_1,\ldots,r_k\}$ and we claim that each triple of maximal density is either tight or spanned.

The proof of LEMMA 1 is based on the following result of Hárs [2]:

LEMMA 2. Let **a**, **b**, **c**, α, β, γ and **A** denote the sides, the opposite angles and the area of a triangle. For given positive weights **u**, **v**, and **w** we consider the weighted *angle-density*

$$\vartheta = \frac{u.\alpha + v.\beta + w.\gamma}{A}.$$

For fixed **a**, **b**, **u**, **v** and **w** the function $\vartheta(\gamma)$ is strictly quasiconvex in $(0, \pi)$, i.e. for any given interval $0 < \gamma_1 < \gamma < \gamma_2 < \pi$ $\vartheta(\gamma)$ attains its maximum only at one or both ends of the interval.

Proof of LEMMA 1.

As the density in a large triangle is small, when looking for the densest arrangement it is enough to consider normal triangles of restricted size. However, the set of normal triangles of sidelength not greater than K is compact, thus the existence of a triangle of maximal density follows easily. Therefore, it is sufficient to show that a normal triangle that is neither tight nor spanned is not one of greatest density.

We consider a normal triangle that is neither tight nor spanned and distinguish two cases.

Case 1. No circle is tangent to the opposite side of the triangle (consequently it is nor spamnned) and there are two circles, say the first and the second that are not tangent (thus it is not tight either). Let us apply LEMMA 2 using the weights

$$u = \rho_1^2.\omega(\rho_1)/2,$$
$$v = \rho_2^2.\omega(\rho_2)/2,$$
$$w = \rho_3^2.\omega(\rho_3)/2,$$

where ρ_j denotes the actual values of the radii (j=1,2,3). (By this choice the weighted angle-density and the weighted density of the circles coincide). The role of γ_1 and γ_2 in Lemma 2 will be plaied by those values of angle γ – the angle opposite to side c – for which the triangle stops being normal, or, with other words, where a further touching occurs (Fig. 2). According to LEMMA 2 the density can not attain its maximum for an angle γ lying strictly between γ_1 and γ_2 thus the triangle in question is nor extremal.

Case 2. One circle, say the third one, is tangent to the opposite side O_1O_2 of the triangle (thus it is not tight), however it does not touch both of the other circles, say the first and the third are not tangent, (therefore it is not spanned). Let us reflect the triangle in straight line O_1O_2 and denote the mirror image of O_3 by O'_3 (Fig 3). Clearly, both isosceles triangles $O_1O_3O'_3$ and $O_2O_3O'_3$ are normal, and, for the densities σ_0, σ_1, σ_2 of the triangles $O_1O_2O_3$, $O_1O_3O'_3$, $O_2O_3O'_3$ it holds

$$A_1\sigma_1 + A_2\sigma_2 = 2.A.\sigma_0 = (A_1 + A_2).\sigma_0,$$

where A_i denotes the area of $O_iO_3O'_3$ (i=1,2). Consequently, σ_0 cannot be the maximum of σ except for $\sigma_1 = \sigma_2 = \sigma_0$. But, since neither triangle is spanned and – according to our assumption in Case 2 – $O_1O_3O'_3$ is not tight either it belongs to Case 1. Hence neither this triangle nor $O_1O_2O_3$ can be of maximal density.

This completes the proof of LEMMA 1.

Remark. Applying the same reflection we used in the discussion of Case 2 it is easy to see that whenever the maximal density is attained by a spanned triangle there is at least one tight triangle of the same density. Consequently, to find the maximal density for a given set of radii $\rho_i \in \mathcal{R}$ and weights $\omega(\rho_i)$ it is enough to compare the densities for the $k+2\binom{k}{2}+\binom{k}{3}$ tight triangles.

4. The THEOREM.

The proofs of the solidity of certain packings can be based on the following general

THEOREM. A packing of circles of radius $r_1, r_2, \dots, r_k$ is solid if

(i) The packing can be decomposed into tight triangles. The actual types of triangles used in this decomposition will be called *tile* triangles.

(ii) Positive weights $\omega(r_i)$ can be assigned to the circles of radius r_i $(i=1,2,\dots,k)$ in such a way that all tile triangles have equal weighted density while the density in any other tight triangle is smaller.

(iii) The union **U** of an arbitrary finite set of triangles of the decomposition can be filled (without gaps and without overlapping) by tile triangles generating a packing only in one way - according to the original pattern.

To prepare the proof of the THEOREM we refrase a result of Fejes Tóth and Molnár [3]:

LEMMA 3. Any saturated packing[1] of circles of radius $\geq p > 0$ can be decomposed into normal triangles - even so that each segment connecting the centers of tangent circles is a side of a triangle of the decomposition[2].

---

[1] A packing of circles of radius p is called *saturated* if there is no room left for a further circle of radius p without overlapping.

[2] This formulation of the result is rather a corollary of their method than the exact citing of a statement in the paper.

Proof of the THEOREM. Let P be a packing for which the assumptions are valid, S an arbitrary finite set of circles of P and S' a rearrangement of these circles that together with the rest P - S of the packing forms a new packing P'. We shall show that P and P' are congruent, i.e. the packing is solid.

Let **U** be the union of a finite set of triangles of the decomposition that covers S and S' as well. Now we define a weighted packing problem for **U**. We consider all sets of circles of radius $r_1, r_2, \ldots, r_k$ that completely lie in **U** and, together with P - S, form a packing and maximize the density of these packings within **U**, when all circles of radius r_i are taken with weight $w(r_i)$ defined in assumption (ii).

Clearly, the original set S provides an extremal solution since **U** can be decomposed into tile triangles each maximizing the density. The contribution of S and S' is the same to the density in **U**, thus the extremality of the corresponding packing implies that P' is saturated. Then - by LEMMA 3 - P' can be decomposed into normal triangles in such a way that the boundary of **U** (consisting of segments each connecting the centers of a pair of touching circles) is not "crossed" by triangles, e.g. **U** is the union of a finite set of these triangles.

From the equality of the contributions mentioned above it also follows that each triangle of this second decomposition of **U** also maximize the density. By LEMMA 1 each of these triangles is either tight or spanned. In fact none of them is spanned, because spanned triangles could occure in P' only in symmetrical pairs implying the existence of touching pairs of circles the centers of which are not connected by a side of triangle. But, this would contradict the basic property of the second decomposition guaranteed by LEMMA 3.

Consequently, all triangles of the second decomposition must be tile triangles. These triangles fill **U** and generate a packing, thus - according to assumption (iii) - the second decomposition coincides with the first one.

This completes the proof of the THEOREM.

References

[1] L.Fejes Tóth, Solid Circle-packings and Circle-coverings, Studia Scientiarum Mathematicarum Hungarica 3 (1968) 401-409.

[2] A.Florian, L.Hárs and J.Molnár, On the ρ-system of circles, Acta Math. Acad, Sci. Hung. **34** (1979), 205-221.

[3] L.Fejes Tóth and J.Molnár, Unterdeckung und Überdeckung der Ebene durch Kreise, Math. Nachrichten, 18 (1958), 236-243.

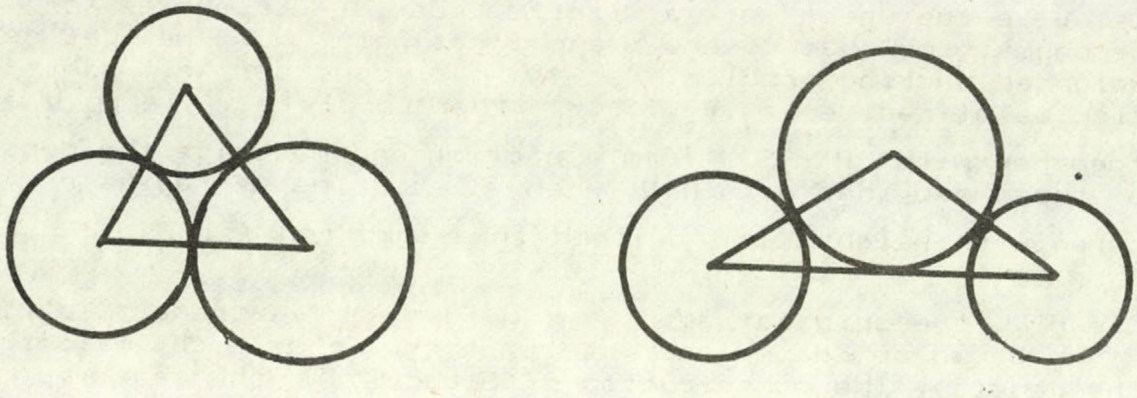


Fig.1

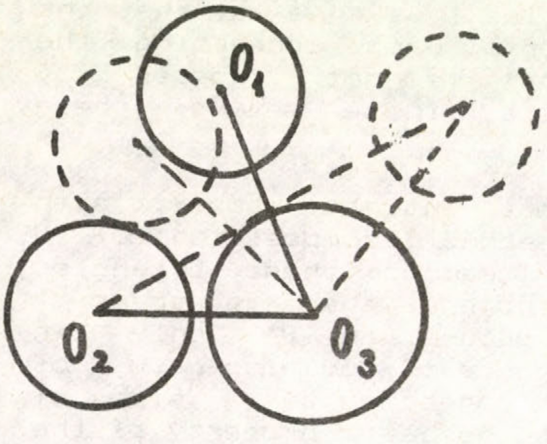


Fig.2

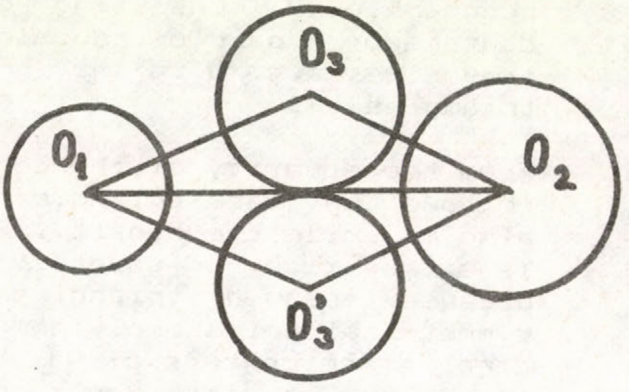


Fig.3

ON THE SOLIDITY OF PACKINGS OF INCONGRUENT CIRCLES I.

A. Heppes

Summary

A packing of convex discs is said to be solid if no finite subset of the discs can be rearranged so as to obtain a packing not congruent to the original one [1].
In the paper a general theorem is proved that contains sufficient conditions for the solidity of circle packings in the Euclidean plane.

INKONGRUENS KITÖLTÉSEK SZOLIDITÁSÁRÓL I.

Heppes Aladár

Összefoglaló

A sik konvex lemezekkel való kitöltését szolidnak nevezzük, ha a lemezek bármely véges részhalmaza csak ugy rendezhető át, hogy az uj kitöltés az eredetivel egybevágó lesz [1]. A cikkben a szerző a szoliditásnak egy elégséges feltételét adja meg.

MTA SZTAKI Közlemények 39/1988 p. 35-52.

A DEDUCTIVE REASONING SYSTEM ON THE BASIS OF A NONMONOTONIC LOGIC

HA HOANG HOP
Institute of Computer Science and Cybernetics
Hanoi - Vietnam

Abstract. This paper presents a deductive reasoning system vs. a set of default theories. Syntactical and semantical aspects of a nonmonotonic logic is considered that provide the background for the deductive reasoning system.

1. Introduction. Nonmonotonicity is the main feature in commonsense reasoning. The statement "Birds fly" is usually given to explain the nonmonotonicity. McDermott and Doyle /1980/ outlines an approach to modeling nonmonotonic reasoning system, McDermott/1982/, Reiter/1980/, Reiter and Crisuolo/1981/, Moore/1983/, Lukaszewicz/1983/ are of much interests in that direction. Various interpretations were made, each gave a specific semantics for a deductive reasoning system. Therefore, it turns out that nonmonotonic logic should be context-sensitive - the set of beliefs of a theory deperds on the determination of a set of axioms for this theory. This paper presents a compromised approach which simultaneously aims to investigate proof-theoretic and model-theoretic aspects of a nonmonotonic logic - modal operators M, L are combined in a single framework of S5-nonmonotonic logic together with a set of default theories. The main intuition is the restriction on the set of needed assumptions when specifying nonmonotonic theorems for a theory. The Computational basis for this deductive system is fixed point properties of an algebraic operator that defines a default theory.

2. Syntactical considerations

Default theories are treated within the framework of propositional language for simplicity sake, after introducing a set of logical axioms and two monotonic inference rules, the nonmonotonic theorems are recognised by terms of modal operators.

2.1. Concepts, definitions and notations.

Definition 2.1.1. Given a classical propositional language Lang /Mendelson - 1965/ which contains:

- a set of proposition letters,
- the set of connectives: $\wedge$ (and), $\vee$ (or), $\sim$ (not), $\Leftrightarrow$ if and only if, ( ) brackets, $\supset$ implication.

to Lang, we attach a modal M "it is consistent", Lang now is usual modal propositional language.

Definition 2.1.2. A term is a constant symbol, a predicate symbol, or an expression $f(t_1,\ldots,t_n)$, where f is a function symbol and $t_1,\ldots,t_n$ are terms.

An atomic formula is an expression $p(t_1,\ldots t_n)$ where p is a predicate symbol and $t_1,\ldots,t_n$ are terms.

A formula is either:

- a proposition letter,
- an expression $\sim p$, where p is an atomic formula,
- $p \supset q$, where p, q are formulas.

Definition 2.1.3. A formula of the form

$$p \wedge Mq_1 \wedge \ldots \wedge Mq_n \supset r$$

or simply

$$Mq_1 \wedge \ldots \wedge Mq_n \supset r$$

where p, ql, ..., q_n, r belong to the classical propositional calculus is named a default.

Definition 2.1.4. A default theory A is a set of formulas together with a set of non-logical axioms of that theory. Each non-logical axiom either belongs to propositional calculus or is a default.

We attach the second modal operator L to Lang, and in the following, L is interpreted as " It is believed ".

Definition 2.1.5. Let p, q, r be formulas in a default theory A. Logical axioms schemata is defined as follows:

/la1/ $Lp \supset p$

/la2/ $Mp \supset LMp$

/la3/ $L(p \supset q) \supset (Lp \supset Lq)$

/la4/ $(p \supset (q \supset p)$

/la5/ $(p \supset (q \supset r)) \supset ((p \supset q) \supset (p \supset r))$

/la6/ $(\sim q \supset \sim p) \supset ((q \supset p) \supset q)$

Monotonic inference rules:

/mr1/ $p, p \supset q \vdash q$ / modus ponens /

/mr2/ $p \vdash Lp$ / necessitation /

where " $\vdash$ " is understood in an ordinary monotonic sense as provability: let S be a set of formulas of a default theory, if p S is provable from S and instances of /la1/ - /la6/ and by application of mr1 and mr2, we denote $S \vdash p$. If not, $S \nvdash p$.

From McDermott and Doyle /1980/, we have

$$Th(S) = \{p: S \vdash p\}$$

It is easy to see that Th has the monotonicity:

/i/ $A \subseteq Th(A)$

/ii/ Let A, B be two default theories, from $A \subseteq B$ we have $Th(A) \subseteq Th(B)$.

/iii/ $Th(Th(A)) = Th(A)$ /idempotence/

The last property of Th can also be viewed as fixed point equation, stating that the set of theorems monotonically derivable from a default theory is a fixed point of the operator which computes the closure of a set of formulas under the monotonic inference rules.

Definition 2.1.6. Let S be a set of formulas. S is consistent if and only if $S \nvdash p$ for only some $p \in S$. A default theory is consistent if and only if its non-logical axioms are consistent.

The above monotonic structure is identical to S5 modal propositional logic /see Hughes and Cresswell, 1972/. In the logical axiom schemata, /la1/ means that everything believable is true, /la2/ shows that p is unprovable only if it provable only if it is provably unprovable, this assertion is useful in nonmonotonic system, /la3/ describes behaviour of modus ponens: it allows to infer q from p q and p, where modus ponens is activated. The instances of the last three axioms /la4/-/la6/ form the axiomatisation for the sentential calculus.

In the following, we settle up the nonmonotonic structure of our default theories, a set of assumptions is added to a default theory by the usual way

Definition 2.1.7. Let d be a default, a formula of the form

$Mq_1 \wedge \ldots \wedge Mq_n$ or simply Mq

is called an assumption of d, and is denoted M d ..

Definition 2.1.8. Let d be a default. Condition for d, denoted by cond d , is defined as follows

$$\mathrm{cond}(d) = \begin{cases} p & \text{if } d = p \wedge Mq_1 \wedge \ldots \wedge Mq_n \supset r \\ p \vee \sim p & \text{if } d = Mq_1 \wedge \ldots \wedge Mq_n \supset r \end{cases}$$

Comment. We give here the similar definition with the ones in Moore /1983/ about objective /resp. subjective/ inference in which we mixture objective and subjective inferences, but define for mixed inferences a condition /in definition 2.1.8/, this serves for convenience of some forms of proof later.

Definition 2.1.9. Let S be a set of formulas, the set of assumptions for S, denoted as As S,d, is defined as

$$As(S,d) = \begin{cases} \{M(d)\} & \text{if } \mathrm{cond}(d) \in S \text{ and } S \cup \{M(d)\} \text{ is consistent} \\ \emptyset & \text{if otherwise} \end{cases}$$

Definition 2.1.10. The set of assumptions for a default theory A, denoted by As_A S , is defined as:

$$As_A(S) = \bigcup_{d \in \text{def } A} As(S,d)$$

where def A denotes the set of all default of A.

Definition 2.1.11. Let A be a default theory and S be any set of formulas. We define operator NM_A as follows

$$NM_A(S) = Th(A \cup As_A(S))$$

Before giving a definition of the special extension, we consider an example belows to clarify some intuitive idea supporting that definition

Example 2.1.12. Consider the theory

$$A = \{ p \wedge Mq \supset q,\ (p \supset q) \wedge Mp \supset p \}$$

There are two fixed points with respect to NM_A : Th(A) and $Th(A \cup \{Mp, Mq\})$. There exists only one extension for A, which is Th(A) , because we have no reason to believe p or $p \supset q$, so it results in the fact that none of the default of A can be activated. The available way to avoid such situations is that by analogy with the monotonic case, we should treate extensions for a default theory A as minimal fixed points of NM_A. We come to the following definition

Definition 2.1.13. Let A be a default theory. A set S of formulas is called a minimal extension for A if and only if S is a minimal fixed point with respect to NM_A, i.e., S is minimal set of formulas such that

$$S = NM_A(S) = Th(A \cup As_A(S))$$

The above definition naturally leads to the following definition of beliefs.

Definition 2.1.14. Let A be a default theory. The intersection of all minimal extension for A is called the set of beliefs derivable from A and is denoted by TH(A).

We have the following theorem.

Theorem 2.1.15. There exists a minimal extension for every default theory A.

Proof. **In the case the default theory A is inconsistent** it is clear that the set of all formulas becomes the only minimal extension for A. With this, now on we may suppose that A is consistent. Our treatment now is to build up a minimal extension for A.

Consider an arbitrary sequence of defaults of A: (d_j). From this sequence we define a sequence (S_i) by the following manner

Put

$$S_1 = Th(A)$$

From a given S_i we define

$$S_i^1 = S_i$$

$$S_i^{J+1} = S_i^j \cup As(S_i^j, d_j)$$

Put

$$S = \bigcup_{i=1}^{\infty} S_i$$

It is easy to see that $S_i^1 \subseteq S_i^2 \subseteq \ldots$

We prove that S is a minimal extension for A, i.e., S is minimal set of formulas and

$$S = Th\left(A \cup As_A(S)\right)$$

S is consistent by induction on i, and also by induction on i, we have

$$S_i \subseteq Th(A \cup As_A(S))$$

which immediately leads to

$$S \subseteq Th(A \cup As_A(S)) \qquad (1)$$

Let $p \in A \cup As_A(S)$. With some $d_k \in def(A)$, we have $p \in As(S, d_k)$. By definition 2.1.9 we have $cond(d_k) \in S$ and $S \cup \{p\}$ is consistent. It implies that for some natural m, $cond(d_m) \in S_m$, furthermore, we have $cond(d_k) \in S_m^k$ because $S_m \subseteq S_m^k$. By the construction of S, we have $S_m^k \subseteq S$. Hence $S_m^k \cup \{p\}$ is consistent. From here we have

$$p \in S_m^{k+1} \subseteq S_{m+1} \subseteq S$$

It implies that

$$A \cup As_A(S) \subseteq S \qquad (2)$$

By definition of Th, we have $S \subseteq Th(S)$.
Let $p \in S$, with $S = \bigcup_{i=1}^{\infty} S_i$, thus

$$\bigcup_{i=1}^{\infty} S \vdash p$$

because $S_o \subseteq S_1 \subseteq \ldots$ and for some natural m, we get $S_m \vdash p$. It implies that $p \in Th(S_m) \subseteq S_{m+1}$.

Altogether we get $p \in S$. So

$$S = Th(S) \qquad (3)$$

From (1), (2), (3) we obtain

$$S = Th(A \cup As_A(S)) \qquad (4)$$

In the rest, we show that S is minimal fixed point.

Suppose that there is a fixed point S_x such that $S_x \subseteq S$. We have $S_i \subseteq S_x$ by the result of induction on i, it implies that $S \subseteq S_x$. Thus $S = S_x$.

This completes the proof of the theorem 2.1.15.

Example 2.1.16./Reiter, Criscuolo, 1981/.

Consider theory $C = \{ p \wedge Mr \supset r,\ q \wedge M(\sim p \wedge \sim r) \supset r \}$. For this theory, we have three possibilities: if p is given, then it is consistent to infer r; if q is given, then it is consistent to infer $\sim r$; if p, q are given simultaneously, it is consistent to infer r. Suppose q is is given, we then have theory $C_q = \{A \cup \{q\}\}$ and its extension $Th(C_q \cup \{M(\sim p \wedge \sim r)\})$. If we add the assumption Mr to A, we then get two extensions for C, which are: $Th(C_q \cup \{M(\sim p \wedge \sim r)\})$ and $Th(C_q \cup \{Mr\})$. The second one contains the formula of the form $p \supset r$ that contradicts to the given conditions, thus we can not accept it in reality.

This example put forwards the fact that when considering a default theory, it is strictly necessary to give attention to those assumptions which are needed for drawing available conclusions.

Notation 2.1.17. The set of needed assumptions for a default theory A is denoted as NA A and

$$NA(A) = \bigcup_{d \in def\ A} \{M(d)\}$$

3. Semantical considerations.

3.1.definitions

Definition 3.1.1. Model is a tuple $M = \langle W, f \rangle$ where W is a nonempty set of possible worlds and f is a function from the set of all proposition letters of Lang to 2^W.

Definition 3.1.2. Let $p \in M$. The truth value for p with respect to $w \in W$, denoted by $u(w,p)$,is defined by mapping v:

$$v: \; W \times S \longrightarrow \{0,1\}$$

so that

$$/t1/ \; v(w,a) = \begin{cases} 1 & \text{iff } w \in f(a) \\ 0 & \text{otherwise} \end{cases}$$

where a is an arbitrary proposition letter in Lang.

$$/t2/ \; v(w, \sim p) = 1 - v(w, p)$$

$$/t3/ \; v(w, p \supset q) = \begin{cases} 1 & \text{iff } v(w,p) = 0 \text{ or } v(w,q) = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$/t4/ \; v(w, Mq) = \begin{cases} 1 & \text{iff } v(w_x, p) = 1 \text{ for some } w_x \in W \\ 0 & \text{otherwise} \end{cases}$$

Definition 3.1.3. Let S be a set of formulas, p is true in M, denoted by $M \models p$, if and only if $v(w,p) = 1$ for every $w \in W$.

Definition 3.1.4. Let S be a set of formula. S is true in M if and only if $M \models p$ for every $p \in S$. In this case we call that M is a model for S.

Definition 3.1.5. Let A be default theory. A set $X \subseteq NA(A)$ is called an activation set of a set $Y \subseteq def\ A$ if and only if the following conditions are satisfied:

/act1/ $A \cup X$ is consistent.
/act2/ $Y = \{d \in def(A): cond(d) \wedge M(d) \in Th(A \cup X)\}$.
/act3/ if $p \in X$, then $p = M(d)$ for some $d \in Y$.
/act4/ for every $d_1 \in def(A) - Y$
$cond(d_1) \in Th(A \cup X)$ or
$Th(A \cup X) \cup \{M(d_1)\}$ is inconsistent.

Definition 3.1.6. A set $X \subseteq NA\ A$ is called a minimal activation set of a set $Y \subseteq def\ A$ if and only if the following conditions are fulfiled:

/ma1/ X is an activation set of Y by definition 3.1.5
/ma2/ There is no activation set of any $Y_1 \subsetneq X$

In the case $Y \subseteq def\ A$ satisfies /ma1/ and /ma2/, we call Y minimally activable.

Definition 3.1.7. Let M be a model for a default theory A. M is called a minimal model for A if and only if M is a model for a minimal activation set $X \subseteq NA\ A$ of a set $Y \subseteq def\ A$.

3.2. Some results.

Theorem 3.2.1. Let A be default theory and suppose that $X \subseteq NA(A)$ is a minimal activation set of a set $Y \subseteq def(A)$. Then $Th(A \cup X)$ is a minimal fixed point with respect to operator NM_A.

Proof. By definition 2.1.13, we have to prove that

$$Th(A \cup X) = Th(A \cup As_A(A \cup X)) \qquad (5)$$

Firstly we prove

$$As_A(Th(A \cup X)) \subseteq Th(A \cup X)$$

Let $p \in As_A(Th(A \cup X))$. There is $d \in def\ A$ such that $p = M(d)$ $cond(d) \in Th(A \cup X)$ and $Th(A \cup X) \cup \{M(d)\}$ is consistent. By the theorem's hypothesis X is an activation set of Y, hence by /act4/ we get $p \in Y$. By /act2/ we have moreover $M(d) \in Th(A \cup X)$, because $p = M(d)$, so $p \in Th(A \cup X)$.

We prove now that $X \subseteq As_A(Th(A \cup X))$.

Let $p \in X$. Because X is an activation set of Y and by /act3/, we have $p = M(d)$ for some $d \in Y \subseteq def(A)$. By /act2/, $cond(d) \wedge M(d) \in Th(A \cup X)$. Therefore, $Th(A \cup X) \cup \{M(d)\}$ is consistent by /act1/ and /act2/. Thus $M(d) \in As_A(Th(A \cup X))$. As $p = M(d)$, so $p \in As_A(Th(A \cup X))$ which completes the proof of (5).

Let $Z \subseteq NA(A)$ be a fixed point of NM_A and suppose that Z is consistent. Consider $As_A(Z)$, we have

$$As_A(Z) = \bigcup_{d \in def(A)} As_A(Z, d)$$

by verifying through /act1/ - /act4/ we conclude that $As_A(Z)$ is an activation set of the set

$$\{d \in def(A) : cond(d) \wedge M(d) \in Th(A \cup As_A(Z))\}$$

Suppose that $Z \subseteq Th(A \cup X)$. We have

$$Th(A \cup As_A(Z)) \subseteq Th(A \cup X) \qquad (6)$$

Denote

$$Y_1 = \{d \in def(A) : cond(d) \wedge M(d) \in Th(A \cup As_A(Z))\}$$

$$Y = \{d \in def(A) : cond(d) \wedge M(d) \in Th(A \cup X)\}$$

From (6) we have $Y_1 \subseteq Y$. But Y is also a minimal activation set, so by (3) / in Thorem 2.1.15 / we get $Y = Y_1$.

It is clear that $Z \supseteq Th(A \cup X)$ because from $Y = Y_1$ we can naturally take $Y \subseteq Y_1$.

The Theorem 3.2.1. is proved.

Theorem 3.2.2 /Completeness theorem/ Let A be a default theory and p be any formula. Then $p \in TH(S)$ if and only if p is true in every minimal model for A.

Proof. The belows lemmae immediatey lead to the completeness theorem.

Lemma 3.2.2.1. /McDermott - 1982, pp.39-40/ Let S be a set of formulas and p be an arbitrary formula in S. Then $p \in TH(S)$ if and only if $M \models S$ for every model for S.

Lemma 3.2.2.2. Let A be a default theory ans Z be a fixed point with respect to the operator NM_A. Then every model M for Z is a minimal model for A.

Proof. Let M be a model for Z. It is clear that every model for Z is also a model for A. Because $Z = TH(A \cup As_A(Z))$ M is a model for Z, so M is model for $As_A(Z)$. By (3) /in Theorem 2.1.15 / $As_A(Z)$ is an activation set for

$$Y_1 = \{d \in def(A) : cond(d) \wedge M(d) \in Th(A \cup As_A(Z))\}$$

In the rest, it suffices to prove that M is a model for a set $X \subseteq NA(A)$ which is an minimal activation set of Y.

Suppose that some set $X \subseteq NA(A)$ is an minimal activation set of a set $Y_1 \subseteq Y$.

Aimimg to prove that X is minimal activation set for Y, we show that

$$Th(A \cup X) \subseteq Th(A \cup As_A(Z)) \qquad (7)$$

To prove (7) it is equivalent to prove

$$X \subseteq Th\ (A \cup As_A(Z))$$

Assume that $p \in X$. For some $d \in Y_1$, we have $p = M(d)$ /by /act3/ /, as $Y_1 \subseteq Y$ we have $p \in Y$. Moreover we have $p \in Th\ (A \cup As_A(Z))$, this completes the proff of (7).

From 4 , we have

$$Th\ (A \cup X) = Th\ (A \cup As_A(Z)) \qquad (8)$$

(8) together with $Y_1 \subseteq Y$ implies that $Y_1 = Y$. This finishes the proof of Lemma 3.2.2.2.

Lemma 3.2.2.3. Let A be a default theory and p be any formula. Then p is true in each minimal model for A if and only p belongs to each minimal fixed point with respect to NM_A.

Proof. /If/ By applying Theorem 2.1.15 we immediately fulfil the "If" part.
/Only if/ This part is direct result of application of two Lemmae 3.2.2.1 and 3.2.2.2.

From lemma 3.2.2.3 we have directly the Completeness Theorem.

4. Conclusion. This paper shows a compromised approach to nonmonotonic reasoning system in comparison with those of McDermott, Doyle, Moore, Reiter and Lukaszewicz : we treate simultaneously two modal operators M and L which allows to consider not only in the light of proof-theoretic but also of model-theoretic aspects, furthermore default theories are manipulated here with the intuitive idea that

every time when a theory is activated, the set of assumptions is carefully considered in order to provide plausible conclusions. We shows the context-sensitivity of our system. It should be noted that our nonmonotonic reasoning system is not semi-decisive, so some intuititions and heuristics are used in building this system - definition 3.5.1 and Theorem 3.5.15 are instances. Moreover, well-defined nonmonotonic theorems are derived from each default theory. Our approach, instead of competing the previous ones, is above all the completion of them.

==========

REFERENCES

1. Hintikka J.,/1962/ Knowledge and Belief: an introduction to the logic of two notions. Cornell Univ. Press, New York
2. Kripke S.A.,/1963/ Semantical considerations on Modal Logic, Acta Phil.Fennica, 16,pp.83-94.
3. Kripke S.A.,/1965/ Semantic Analysis of Intuitionist Logic I, in: Crossley J.N. and Dummel M.A.E. /Eds/, Formal Systems and Recursive Functions, North-Holland, Amsterdam, pp.92-130.
4. McDermott D., Doyle J.,/1980/ Nonmonotonic Logic I, Artificial Intelligence 13/1,2/,pp.41-72.
5. McDermott D.,/1982/ Nonmonotonic Logic II: Nonmonotonic Modal Thories, JACM, 29/1/, pp.34-41.
6. Hughes G.E., Cresswell M.J.,/1972/ An Introduction to Modal Logic, Menthuen and Co., London.
7. Mendelson E., /1965/ Introduction to Mathematical Logic, Van Nostrand Reinhold, New York.

8. Moore R.C.,/1983/ Semantic Considerations on Non-monotonic Logic, Proc. of 8th IJCAI-83, Karlsruhe, West Germany, pp.272-279.

9. Rasiowa H., /1974/ An Algebraic Approach to Non-Classical Logics, North-Holland, Amsterdam.

10. Reiter R.,/1980/ A Logic for Default Reasoning, Artificial Intelligence 13/1,2/, pp.81-132.

11. Reiter R., Crisuolo, R.,/1981/ On Interacting Default, Proc. IJCAI-81, Vancouver, pp.81-132.

12. Watanabe S., /1969/ Knowing and Guessing, John Wiley and Sons Inc., New York.

13. Lukaszewicz W., General Approach to Nonmonotonic Logics, Proc of 8th IJCAI-83, Karlsruhe, West Germany, pp.352-354.

A DEDUCTIVE REASONING SYSTEM ON THE BASIS OF A NONMONOTONIC LOGIC

Ha Hoang Hop

Summary

The paper presents a deductive reasoning system, where both syntactical and semantical aspects of a non-monotonic logic are considered.
Non-monotonicity is the main feature in commonsense reasoning. Many approaches to modelling non-monotonicity are known. The author presents a compromised approach which simultaneously aims to investigate proof-theoretic and model-theoretic aspects of non-monotonic logic.

EGY NEM-MONOTON LOGIKÁN ALAPULÓ DEDUKTIV KÖVETKEZTETÉSI RENDSZER

Ha Hoang Hop

Összefoglaló

A cikk egy deduktiv következtetési rendszert mutat be, amely a nem-monoton logikák mind szintaktikai, mind szemantikai aspektusain alapszik. A szerző a nem-monotonitásnak /amely a "józan következtetésnek" fő tulajdonsága/ egy kompromisszumos modelljét mutatja meg, amely a nem-monoton logikák mindkét tárgyalásának /modell-elméleti illetve bizonyitás-elméleti/ aspektusait felhasználja.

MTA SZTAKI Közlemények 39/1988 p. 53-66.

A DEDUCTIVE LANGUAGE FOR THE REPRESENTATION OF INCOMPLETE KNOWLEDGE

HA HOANG HOP

Institute of Computer Science and Cybernetics
Hanoi - Vietnam

Abstract We give here a formal treament of the problem of representing incomplete knowledge by a new approach in which formal knowledge bases are described in terms of abstract interactive operations between them and expert systems /ESs/. We then setle up an original knowledge representation language which allows the explicit representation of nonmonotonicity in the framework of beliefs and knowledges. The correctness of this language is proved.

1. Introduction Our formalisation of RL - the representation language - is started with McDermott and Doyle /1980/, Moore /1983/, Levesque /1984/. Levesque s paper provides much technical inspirations. We generalise Levesque s knowledge representation language by two foundamental aspects: all interactive operations are abstract instead of only as "Tell" and "Ask", and beliefs are more explicitly represented by means of simultaneous two algebraic operators K "It is known" and L "It is believed". The first aspect allows a high level of conceptual modeling, since with two operations "Tell" and "Ask", formal knowledge bases would have the risk of colliding to traditional databases in which well-formed formulae are only considered in part; the second aspect enables a more explicit representation of beliefs, as we explicitly specify in RL the second algebraic operator L - we distinguish belief and knowledge, this distinction seems to be essential when dealing with the Generalised Closed-World Assumption /Reiter -1978, Bossu and Siegel-1985, Przymusinska and Gelfond-1986/

We give an approach to above problem by a new manner: turning reasoning steps into so-called knowledge matrices and then showing the correctness of RL.

2. The knowledge representation language RL

2.1. Concepts, definitions

We buil up RL on the basis of the following sets:

W, KB, F, R

where

. W is a non-empty set of possible worlds,

. KB is a set of preliminary knowledges about W. In the following, KB is treated as an abstract knowledge base.

. F is the set of all formulae of FOL and all rmulae of the form:

/1/ $p \wedge Kq_1 \wedge \ldots \wedge Kq_n \supset r$ or simpler $Kq_1 \wedge \ldots \wedge Kq_n \supset r$

/2/ $p \wedge Lq1 \wedge \ldots \wedge Lq_n \supset r$ or simpler $Lq_1 \wedge \ldots \wedge Lq_n \supset r$

where p, q_i (i=1,...,n), r belong to FOL. K is quoted as "it is known" and L is as "it is believed".

. R is a set of specific representations so that each $r \in R$ transforms an arbitrary pair (w,d) where $w \in W$ and $d \in F$ into an element of

C= { yes, no, unknown, known, believed }

Definition 2.1.1. Axioms for FOL proof theory:

/a1/ $p \supset (q \supset r)$.

/a2/ $(p \supset (q \supset r)) \supset ((p \supset q) \supset r$.

/a3/ $(\sim q \supset \sim p) \supset ((\sim q \supset \sim p) \supset q$.

/a4/ $\forall x (p \supset q) \supset (\forall x p \supset \forall x q)$.

/a5/ $\forall x p \supset p^x_t$.

Axiom of Equality:

/ae/ $(i=i) \wedge (i \neq j)$ for all distinct i, j p, q, r are formulae of FOL, x is free variable, t is closed term and i, j are indexes.

Comment 2.1.2. Since KB is defined as preliminary knowledges from W, it puts equivalence to that KB is incomplete.

With the algebraic operator K, we have Levesque s query language in which a formula of the form Kp is read as

The KB knows p /ak/

or

KB $\Vdash$ p

where $\Vdash$ is a specified provability relation and p is any formula of FOL possibly containing K′s.

The second algebraic operator L is as

The KB believes p /ab/

or

KB $\Vvdash$ p

where $\Vvdash$ is a new form of provability /or query evaluation/

While the first query evaluation $\Vdash$ is understood as usually, the second needs an exact semantics.

2.2. Semantics for RL

The language RL has all formation rules of FOL and the following two rules:

If p $\in$ RL then Kp $\in$ RL /ak/
If p $\in$ RL then Lp $\in$ RL /ab/

So, in F there is three kinds of formula:

/i/ $p \wedge q$
/ii/ $K(p \wedge q)$
/iii/ $L(p \wedge q)$

/i/ will be true or false depending on the interpretation of predicate symbols.
/ii/ will be true or false depending on KB and on what is known or unknown.
/iii/ will be true or false depending on KB and on what is believed.

Semantics of RL will be depended on the set of possible worlds W. We use here Kripke s interpretation. Kripke /1963/ uses the concept of possible worlds to create a formal semantics for modal logic. Later, mathematical logicians, e.g., Chang and Keisler /1973/, Kalish et al./1980/, equate the concept of possible worlds with the model for a formal language of FOL. In another development, linguists and philosophical logicians,e.g., Cresswell /1973/, Rescher /1975/, seem to regard possible worlds more broadly, as a kind of Gestalt experiments, not limited by the vocabulary of the language of FOL or any others. Our usage of possible worlds here will be more on mathematical side,i.e., that a possible world is an alternative model, but the philosophical side remains valid.

Definition 2.2.1. Kp is true in KB iff p is true in every possible world of W.
Definition 2.2.2. Lp is true in KB iff p is true in every possible world of W.

In agreement with McDermott /1982/ s argument, we need two inference rules: modus ponens and necessitation. We come to the axiom schemata for RL.

Definition 2.2.3. Axiom Schemata for The Knowledge Representation Language RL

. The axioms of FOL

. Kp where p is an axiom of FOL
. Lp where p is an axiom of FOL
. $K(p \supset q) \supset (Kp \supset Kq)$
. $(\forall x) Kp \supset K(\forall x) p$
. $L(p \supset q) \supset (Lp \supset Lq)$
. $(\forall x) Lp \supset L(\forall x) p$
. $p \equiv Kp$ where p is pure.

/ a formula is pure when it is known /

Definition 2.2.4. Monotonic inference rules for RL

. $p, p \supset q \vdash q$ /modus ponens/
. $p \vdash Kp$ /K-necessitation/
. $p \vdash Lp$ /L-necessitation/

Comment 2.2.5. On the basis of the notion of possible worlds, K and L are treated in an unified way: definitions 2.2.1 and 2.2.2 have pointed out this unification, thus in the above definition, we may use $\vdash$ instead of $\Vdash$ and $\models$ without confusions. The semantics for RL is both sound and complete since RL is both sound and complete with respect to Levesque /1984/ and Hop /1987/. Due to the limited space, we do not quote the proofs here.

Theorem 2.2.6. RL is both sound and complete

3. The correctness of RL

In this section, we consider more concretely what RL will be with abstract interactive oprations. We define exactly interactive operations by terms of specific representation, then gradually, the so-called knowledge matrix is constructed that serves for the proof of the correctness of RL - this proof in turns, is strongly not only supports for theorem 2.2.6, but opens new possibilities for further investigations.

3.1. Definitions

Definition 3.1.1. $r \in R$ is defined as follows:

$$r: KB \times RL \longrightarrow \{\text{yes, no, known, unknown, believed}\}$$

so that

$$r(k, p) = \begin{cases} \text{yes, if } Kp \text{ true on } k \\ \text{no, if } K\ \ p \text{ true on } k \\ \text{believed, if } Lp \text{ true on } k \\ \text{known, if } p \text{ is pure} \\ \text{unknown, otherwise} \end{cases}$$

Definition 3.1.2. The quantification of R:

$$Q\big(r(k,p)\big) \rightarrow \{0, \Delta_1, \Delta_2, \Delta_3, 1\}$$

so that

$$Q(r) = \begin{cases} 1 & \text{if } r(k,p) = \text{yes} \\ 0 & \text{if } r(k,p) = \text{no} \\ \Delta_1 & \text{if } r(k,p) = \text{believed} \\ \Delta_2 & \text{if } r(k,p) = \text{known} \\ \Delta_3 & \text{if } r(k,p) = \text{unknown} \end{cases}$$

where $0 \leqslant \Delta_3 < \Delta_2 \leqslant \Delta_1 \leqslant 1$.

Definition 3.1.3. The classification of algebraic/modal formulae:

/i/ if p is formula, then so Kp, Lp.

/ii/ Classification listing for all formulae / with or without algebraic/modal operators /

α	α_1	α_2
$(p \wedge q, 1)$	$(p, 1)$	$(q, 1)$
$(p \vee q, 0)$	$(p, 0)$	$(q, 0)$
$(p \supset q, 0)$	$(p, 1)$	$(q, 0)$
$(\sim p, 1)$	$(p, 1)$	$(q, 1)$
$(\sim p, 0)$	$(p, 0)$	$(q, 0)$

β	β_1	β_2
$(p \wedge q, 0)$	$(p, 0)$	$(q, 0)$
$(p \vee q, 1)$	$(p, 1)$	$(q, 1)$
$(p \supset q, 1)$	$(p, 0)$	$(q, 1)$

υ	υ_0
Kp, 1	p, 1
Lp, 0	p, 0

π	π_0
Kp, 0	p, 0
Lp, 1	p, 1

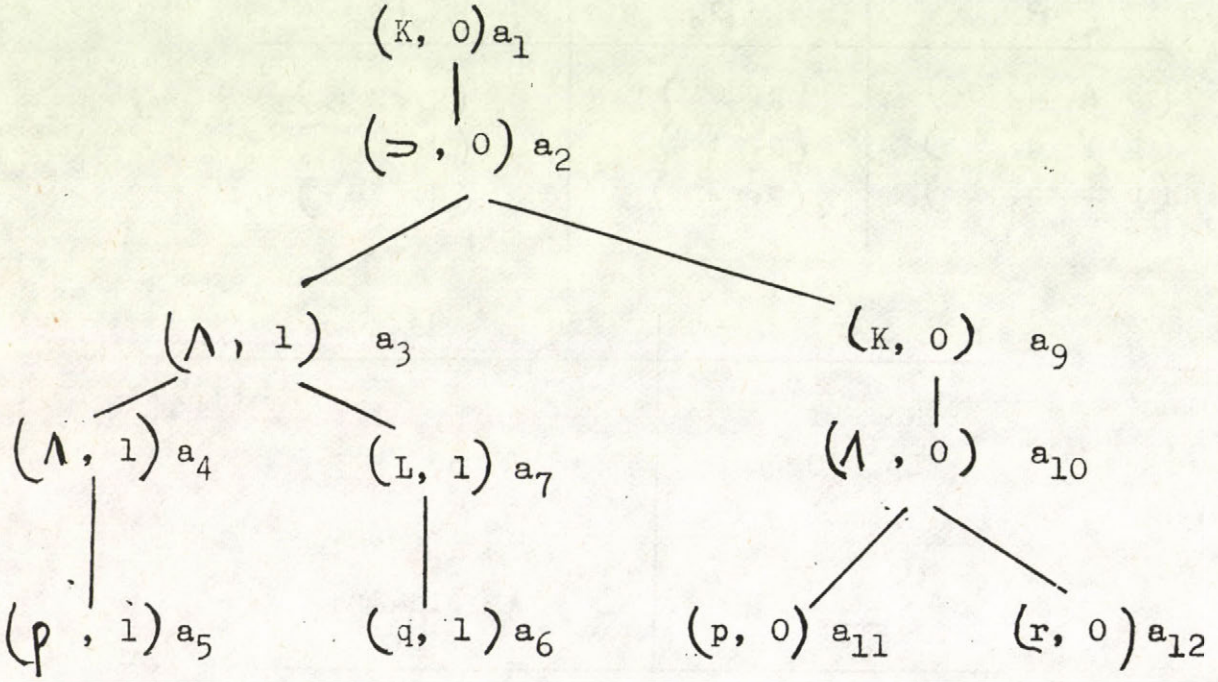
where $\alpha, \alpha_1, \beta, \beta_1, \beta_2, \alpha_2$ --- are used to denote signed formulae and their components of respective types. This classification is a modification of Hop /1986/ for temporal case, however, we give here only a classificatio which relates to classical quantification with two values 0 and 1 - this treatment is enough for the matrix representation that we will present belows.

Definition 3.1.4. The formula tree for a signed modal formula is a formation tree containing additional information as to the polarity of its atomic formulae.

Example 3.1.5. The formula tree for the formula

$$(K(p \wedge q) \wedge Lp \supset L(p \wedge r), 0)$$

is



the label at each node is index of the path in the tree, which provides the following matrix representation of formulae

Definition 3.1.6. A path through a formula tree is a subset of its formula tree. Denote these paths s and t, and s u for path s with an occurence of the label u .

A path through a formula tree is called atomic iff for every node k in s, either

/a/ k is a label by an atomic formula, or

/b/ k is a v in the classification tables.

The atomic paths through an quantified formula is detected by wringting components of an α-type side by side and the components of β-type one above the other to form a nested matrix.

Example 3.1.7. The matrix representation of the formula in example 3.1.5:

$$-- K - \left(-- \left(-p \wedge q- \right) --- \left(-Lr- \right) \right) \supset --L-- \left(\begin{pmatrix} p \\ \wedge \\ r- \end{pmatrix} -- \begin{pmatrix} p \\ \wedge \\ -r- \end{pmatrix} \right) ----$$

one of four atomic paths is a dotted line in the figure.

With $r \in R$, a pair k,p is transformed into an value in the set $\{ 0, 1, \Delta_1, \Delta_2, \Delta_3 \}$. It in turn, is transformed by quantification Q into one of the categories in classification tables, and at last, comes into the form of matrix representation above. We have the following definition.

Definition 3.1.8. Through r, Q and matrix representation, we obtain matrix $\| x_{ij} \|_{qxm}$ forming by elements: 0, 1, Δ_1, Δ_2, Δ_3, where q is cardinality /if any/ of p in RL and m is cardinality of KB /if any/. Thus, in the simplest case, KB can be represented as a matrix, and in general, it can be represented as matrices. We call this matrix /or matrices/ knowledge matrix /or matrices /.

We obtain the Theorem belows which points out the existence of specific representation for every interactive operation between KB and expert systems. This theorem actually shows the correctness of every operation between KB and their abstract expert systems.

Theorem 3.1.9. For every $r \in R$, there exists two other specific representations $r_1, r_2 \in R$ so that $r = r_1 \circ r_2$, where "o" is a consequent application of r_1 and r_2, i.e., a matrix multiply, and

$$r_1\ (k,p) = \|k_{ij}\|_{q \times m}\ ,\ r_2(\|k_{ij}\|_{q \times m}) = \|x_{ij}\|_{qxm}$$

$$x_{ij} \in \{0,\ 1,\ \Delta_1,\ \Delta_2,\ \Delta_3\}$$

Proof. Let r_3 be a specific representation that transforms $\|x_{ij}\|_{qxm}$ to a quantified matrix $\|y_{ij}\|_{qxm}$ / this situation is realistic since we can put $x_{ij} = y_{ij}$ if $x_{ij} \in \{0, 1\}$, and $y_{ij} = 1/2$ if $x_{ij} \in \{\Delta_1, \Delta_2, \Delta_3\}$ /. Thus by r_3^{-1} we denote the reverse representation from $\|y_{ij}\|_{qxm}$ to $\|x_{ij}\|_{qxm}$.

Put $r_1 = r_1 \circ r_3,\ r_2 = r_3^{-1}$ we have

$$r = r_1 \circ r_2 = r_1 \circ r_3 \circ r_3^{-1} = r.$$

The Theorem is proved.

4. Conclusion. With RL, we can resolve the following problems:

. The interaction between an abstract expert system with an incomplete knowledge base by non-monotonic manner,

. Formal semantics, a Kripke semantics-based consideration is explicitly investigated,

. For the first time, we propose the matrix representation for the problem of constructing up an knowledge representation language, this treatment shows that even with extended power provided by RL, knowledge of an incomplete knowledge base is still representable at symbol level, and even at number level by means of two modal operators in the framework of matrix representation. This means that whenever we want to have the "returns" to traditional databases, it will be quite possible.

Recently almost all contributions to the field of knowledge representation are deductive, i.e., are based on various deductive reasoning models. It means that the logical lattice of deduction models are supposed closed. The closure property is presented inTheorem 3.1.9. Closures are abstractly and generally investigated by Khang /1978/, thus it seems to enable an unified approach to pattern recognition and knowledge representation simultaneously.

References

1. Chang, C., and Keisler, H.J. /1973/. Model Theory. North-Holland, Amsterdam.
2. Cresswell, M.J. /1973/. Logics and Languages. Methuen and Co., London.
3. Hintikka, L. /1962/. Knowledge and Belief: An Introduction to The Logic of Two values. Cornell Univ. Press.
4. Ha Hoang Hop /1986/. An Algebraic Approach to Knowledge Structuring. Proc. of KNVVT on Aotomation of Information Processes by Personal Computers. SzTAKI, Tanumányok, No. 194, pp.73-81.
5. Ha Hoang Hop /1987/ Formal Logic-based Approach to Knowledge Representation. Thesis. MTA-KFKI. Submitted to Hungarian Academy of Sciences.
6. Kalish, D., Montague , R., and Mar, G., /1980. Logic - Technics for Formal Reasoning, Harcourt, New York.
7. Bach Hung Khang /1978/. Algebraic Closures in Pattern Recognition. Thesis of Doctor Degree. The Academy of Sciences of The USSR, 350 p.
8. Levesque, H. /1984/. Foundations of Functional Approach to Knowledge Representation, Artificial Intelligence 23, pp.155-212.
9. McDermott, D,and Doyle, J. /19 80/. Nonmonotonic Logic I, Artificial Intelligence 13/1, 2/, pp.34-41.

10. Mendelson, E. /1965/. Introduction to Mathematical Logic. Van Nostrand Reinhold, New York.

11. Moore, R.C.,/1983/. Semantic Considerations on Non-Monotonic Logic, Proc. of 8th IJCAI-83, Karsruhe, West Germany, pp.272-279.

12. Rasiowa, H, and Sikorski, R. /1970/. Mathematics of Metamathematics, PWN, Warsaw.

13. Reiter, R. On Closed-World Databases, in: Gallaire, H., and Minker, J. /Eds/, Logics and Databases, Plenum, New York, pp.55-76.

14. Rescher, N. /1975/. A Theory of Possibilities - A Constructivistic and Conceptualistic Account of Pössible Individuals and Possible Worlds. Pittsburg Univ. Press.

15. Gelfond, M., and Przymusinska, H. /1986/. Negation as Failure: Careful Closure Procedure. Artificial Intelligence, 24, pp.13-63.

16. Bossu, G., and Siegel, P. /1985/. Saturation, Nonmonotonic and the Closed-World Assumption. Artificial Intelligence, 24, pp.13-63.

A DEDUCTIVE LANGUAGE FOR THE REPRESENTATION OF INCOMPLETE KNOWLEDGE

Ha Hoang Hop

Summary

In the paper a formal treatment of the problem of representing incomplete knowledge is given. By this a new approach to formal knowledge bases is described in terms of abstract interactive operations between them and expert systems (ES). An original knowledge representation language is introduced, that allows the explicite representation of non-monotonicity in the framework of beliefs and knowledges. In the paper the correctness of this language is also proved.

EGY DEDUKTIV NYELV A NEM TELJES TUDÁS REPREZENTÁLÁSÁRA

Ha Hoang Hop

Összefoglaló

A cikk a nem teljes tudás reprezentálásának egy uj formális tárgyalását adja. Ez a tárgyalás a formális tudásbázisok és a szakértő rendszerek közötti absztrakt interaktiv leképezésén alapszik. A szerző bevezet egy eredeti tudás-reprezentáló nyelvet, amely a "nem-monotonitásnak" egy explicit reprezentálását is lehetővé teszi, éspedig a tudás illetve vélemény fogalmak keretében. A szerző a nyelv korrektségét is bizonyitja.

MTA SZTAKI Közlemények 39/1988 p. 67-82.

A KNOWLEDGE-BASED APPROACH TO SOFTWARE REQUIREMENTS SPECIFICATION

A.M. KAROL

Tsentprogrammsystem,
Kalinin, USSR

1. INTRODUCTION

During the last decade requirements development stage has gained the status of probably the most critical stage in the whole process of fairly complex software development. As a matter of experience [1], late discovery of errors made at this stage leads to serious alterations of the system and significantly affects the time and development cost.

The basic goal of the stage is a clear and complete, modifiable description of what the conceived software system should perform, the description which is unambiguously understood by users, experts and developers.

Among multiple aides proposed for achieving this goal, the essential part consists of those specification methods and systems, in which required information, recorded in some formal or semiformal language, is stepwise added to a common design database and then, if required, can be subjected to automatic semantic checking, modified or printed in the form of structured reports, constituting the target system documentation.

The most known example of the systems of this kind is the PSL/PSA system described in details in literature [2], which was the first to define the above mentioned scheme and served as a starting-point of many approaches [3] that promoted not only theoretical comprehension of a problem, but accumulation of practical experience in application of computer-aided systems for specifying the early stages of software development.

Nevertheless, it should be recognized that nowadays, the application of such aides in broad programming practice has episodical nature. There are many reasons for that. Without going into details we will name here only two reasons of considerable importance in our opinion. The first one is complexity and unwieldiness of the requirements specification languages and systems, and related with them difficulties in learning and efficient usage. The second reason consists of the fact that such systems more frequently do not allow us to check dynamic aspects of the system being developed.

This paper describes the SAM system (Specification, Analysis, Modelling) [4], which is a knowledge base that can be filled with knowledge, and it is intended for development of software requirements and allows smooth transition to more detailed specifications. In particular, SAM tries to overcome, to some extent, the two above mentioned deficiencies characteristic for many specification systems. The former at the expense of the user friendly interface that involves menus, windows, and a user defined specification language; the latter at the expense of the application of the Prolog language [5] as a tool for rapid prototyping [6]. In many respects SAM continues the line "PSL/PSA - SDLA [7,8,9] - SAD [10]". For example, SAM does not fix previously defined technology of usage, and can serve as a tool supporting different technologies. SAM is based on the ideas of conceptual database modelling and logic programming elements.

2. THE SAM BACKGROUND

Modern understanding of the requirements definition problem assumes, if possible, a more complete description not only of a system itself, but the surrounding world, in which the system will function interacting with this world [11]. The main motivation for such an extended interpretation lies in the fact that requirements can be formulated satisfactorily only when the system application area is understood well [12], and thus some "sum of knowledge" about the target system and the part of the real world it is related to, can be fixed. Similar trend, concerning the knowledge representation techniques, can be observed in database design - the field in which a conceptual model is used to describe a conceptual scheme (universe of discourse) - of the part of the real world which defines the database contents [12].

The SAM system model allows for the presented considerations. It originates from the SDLA system semantic data model, and the Horn logic implemented in the Prolog language. Such a combination in this case is not merely a mechanical connection, some features of the modified extended relational data model of the SDLA system and Prolog, which itself one can treat as generalization of relational databases, do it natural.

In order to make the paper self-contained we briefly describe basic features of the SDLA system developed at the Computer and Automation Institute of the Hungarian Academy of Sciences in co-operation with the ISDOS project of the Michigan University.

The SDLA system implements a meta language specification approach [13], i.e. it supports not one specific language (like, for example, PSL in the PSL/PSA system), but provides each user with the capability to define and expand his own problem-oriented, easily understandable language at the meta level. Later this language can be used at the second level - at the level of target system description, and also when database queries are formulated.

Besides a concrete language syntax, a conceptual scheme (a set of types of objects or "concepts" in the SDLA terminology) and some semantic constraints are defined at the meta level. At the description level the system checks correspondence of input objects (concept instances) to appropriate types and carring-out of the specified constraints.

The concept is described by its attributes:

concept name_of_concept (sel1:attr1,...,selN:attrN),

where seli (selector) is a name of an i-attribute. Attributes should also be defined as concepts; recursion is allowed. Definition is also correct in the case when the part in parentheses does not exist. There is an important mechanism for refining the concepts - a scheme of types and subtypes. Let's mark one more feature in the SDLA database, which is interesting as such and from the point of view of its syntactic generalization and semantics redefinition in SAM. If at the meta level two concepts c1 and c2 (such that a nonempty set of attributes of the concept c2 is an own subset of a set of attributes of the concept c1) are connected by the relationship

$$c1(sel1{:}attr1,\dots,selN{:}attrN) \text{ implies } c2(seli_1{:}attri_1,\dots,seli_k{:}attri_k) \quad (k < N),$$

then at the object level the creation of an instance of the concept c1 automatically causes the creation of an instance of the concept c2.

3. BASIC FEATURES OF SAM.

From a very general point of view the objective goal of SAM is a system description and analysis, just like the goal of the SDLA system. However, SAM has a number of fundamental differences expanding its functional capabilities, and reflecting a specific character of the SAM implementation on a personal computer.

SAM is based on a knowledge base, in which relevant information about a subject domain (general knowledge) and a system being developed (computer artifacts) is accumulated. Interaction with the knowledge base is performed at two levels: a meta level and an object level. A knowledge unit specified in the SAM system input language at the meta level, like in SDLA, is referred to as a concept. A concept is a carrier of common features of objects - knowledge units of the object level. On a par with the term "concept" we shall use as synonims also the terms "type", "object type". Objects are described by using key words, which are fixed the moment the concept, to which these objects belong to, is defined. The key words make up so-called forms (patterns, sentences), for each of which, moreover, the context and the name description in which it can be used, is specified. For each concept one or more forms can be specified.

Concepts and objects may represent simple and complex, real and abstract, static and dynamic phenomena, entities, and relations. People, enterprises, problems, machines, computers, DBMSs, transactions within DBMS, technologies, numbers, module invocation, subsystem membership etc. are some examples.

A person describing a concept gives a suitable name to the concept and specifies, if necessary, a list of its named components (attributes). Components, in their turn, are defined as concepts. Another possible way of defining the concepts is to add new components to the name of previously described concept, and it is referred to as concept refinement or specialization [15]. This way is implemented by the built-in concept IS.

Built-in concepts are concepts that have been defined inside the system. Among them are system data types and system operations. System data types are integers, reals, and texts. Objects of these types are also considered known in the system. The system operations include equality and unequality predicates defined for the objects of the same type, comparison predicates >, <, >=, =< defined on system data types, arithmetic operations defined on numbers. Other built-in concepts will be considered below.

The most essential difference between the SAM system and the SDLA system is in the fact, that it is possible to state rules connecting the concepts described at the meta level in SAM. The rules are written roughly in the form as it is done in the Prolog language. Formally, the record

```
c0(...) if c1(...) and  ... and cN(...).        (N >= 0)
```

means, that for all values (objects of corresponding concept attributes) of the variables it contains, the concept c0(...) holds when c1(...) and ... and cN(...) hold. Another valid variant of a rules record employes forms of the concepts c0, c1, ..., cN.

If in the rule there is a concept with one argument (variable or constant), which is preceded by the # symbol, this indicates that objects of the concept itself are taken into account, rather than objects of its attributes. That way is sometimes more convenient, but the first one gives better flexibility. On the whole, this correspondence between composite objects and tuples of their components plays an important role in building of the SAM system.

In the right part of the rule it's also possible to use the logical connectives or, not. Logical operations priorities are usual, however they can be changed by parentheses.

Rules allow us to represent semantics of interconnections between concepts. Like concepts and objects, rules can be added, deleted and edited by a designer. Each rule has its name.

Rule interpretations can be various. Rules with an identical interpretation are integrated into named groups. At each point of time some rule groups are marked as "active".

Rules describing subject domain integrity constraints serve as an example of rules integrated into one group (A). Rules representing specific character of a particular task within the framework of one subject domain can constitute another group (B).

If new objects being added at a time can, to a designer's opinion, violate an integrity of a knowledge base being designed, tnen the rule groups A and B should be pointed as active.

Rules specifying deductive capabilities of a knowledge base being created, costitute a specific group. This group is activated in a query phase, and allows us to represent some information intensionally, in terms of logical implication, instead of describing and storing objects explicitly. A sample rule from this group is the rule

grandfather(X,Y) if father(X,Z) and parent(Z,Y).

In this case grandfather(...), father(...), parent(...) are concept names, and variables X,Y,Z can be names of objects from the knowledge base addressing the concept man(...).

In a special mode of the SAM system one can exit to the basic language level, this language is Prolog. In other words, it is allowable to use in the rules built-in predicates of Prolog and new "pseudo-concepts", defined as the Prolog language procedures in this mode. Such a capability allows "to program" new rule groups interpretations, and it is aimed mainly for Prolog language programmers. On the other hand, direct exit to Prolog can be useful for a designer too for purposes of rapid prototyping of fragments of the system being designed [16,17], and for making unrestricted algorithmic processing of the database contents. Such an activity requires familiarity with Prolog. It should be noted that the description style in the SAM system and the programming style in the Prolog language have some similarities. This helps designers learn and utilize Prolog.

The second feature of SAM also partly finds its analog in Prolog - the language in which a sort of the meta level (predicates and rules with variables only) and the object level (facts) are not devided explicitly. When a concept is defined in SAM, a set of values for a component can be made narrower in regard to the specified attribute. This can be achieved by specifying a list of names of objects of the corresponding type, and by specifying a range or a fixed value for system data types. Thus, there is no need to describe a separate concept, and the advantages of type checking are used in full measure. Moreover, such a capability allows us to consider concepts, the objects of which have common characteristics (a special case of an object association [15]).

An important additional feature of the SAM system is the support of multi-aspect description of the system being developed, i.e. "decomposition" of the system description into "orthogonal" components. Each component is an entire system description but it depicts only one subset of its properties [14]. Hierarchical system decompositions made by several (up to 7) properties or aspects, significant from a designer's standpoint are allowed in SAM. For each property designer selects one of the built-in concepts UNDERk (k=1,...,7). For any k UNDERk there is a binary relation, defining a hierarchy by a given property. A sample aspect can be any of the abstraction mechanisms available in SAM (aggregation, specialization, association); a property allowing for different-type meaningful concept "subordination" (amalgamation - plant - shop - team); a feature presenting a sequence of execution of singled out functional elements of the system being designed. However, nonhierarchical descriptions are allowed. A complete description of the system consists of a set of named hierarchical and nonhierarchical descriptions.

The SAM query language is closely related to various types of descriptions. It includes two parts. The first predefined part has the most frequently used "command" queries. The second query part is formed by a designer in the same language, which he used in describing the objects of the system under development.

One more specific feature of the SAM system is processing of undescribed objects and concepts. When concept components are described, an identifier, which is not a name of any defined or built-in concept, can be pointed. The system warns a designer about this and places the pointed identifier into a list of nondefined concepts. The list can be looked through in an appropriate screen window. As long as the nondefined concepts list is not empty, the system limits the designer's activities by warning him that there are nondefined concepts. Similarly, the system maintains the nondefined object name list, by grouping the objects by names of "their future" types.

4. USER INTERFACE

SAM is an interactive system using a multi-window interface and menus. Different knowledge base status information can be displayed in windows: in one window a list of concepts available in a given moment is displayed, in another - a list of objects for some concepts, in the third - a list of names for rule groups etc. A typical system session is as follows. At the beginning of an operation the system is in a main menu state. A user can select an operation with concepts, objects (or descriptions), rules, can query the knowledge base, familiarize oneself with the short system operation instruction, or logout the system.

Suppose, for example, a user wants to create a new concept. He should mark the corresponding line CONCEPTS in the main menu. After this a new menu with a list of actions that can be performed on the concept will appear on the screen, and in the right-hand corner of the screen a window with a list of the defined concepts will be displayed. When the line CREATE is selected in the last menu, a pattern for the newly created concept will appear on the screen, and the cursor is placed into the position, from which the new concept name should start. Then if the name is entered correctly, the cursor is moved to the beginning of the next field to be filled. This field is the selector field of the first component of the concept being created. If a user has typed the name of the previously created concept, the system will inform us about this in the bottom horizontal window (intended for messages), and will return the cursor to the first concept name field position, and so on.

The object description is reduced to filling the empty places in an appropriate form. Up to this moment by pressing the functional key the system allows us to look over the list of forms permissible in a given place of the given description.

5. SYSTEM IMPLEMENTATION

As it has already been mentioned, the SAM system is implemented in the Prolog logic programming language. Declarative and procedural semantics of the Prolog programs allows us to combine declarative and procedural knowledge representation and facilitates the implementation of the SAM system knowledge representation language. Moreover, as long as a Prolog program can dinamically modify itself (add and delete clauses during execution), manipulation of concepts, objects and rules is facilitated in SAM. Switching from the meta level to the object level and vice versa can be done at any time, and it does not depend upon scheme modification at the meta level.

The SAM system architecture includes some components of expert systems - a knowledge base, an inference engine, a user interface. Presentation about the knowledge base and user interface is given above. The purpose of the inference engine is to provide switching from one rule group into another, and to perform the required search and transformation in the knowledge base.

The system is being developed for IBM /PC/XT/AT and compatible under the MS DOS operating system.

6. CONCLUSION

We have considered an approach towards software requirements development, which is based on knowledge base technology. The essence of the approach is to combine the capabilities of the PSL/PSA trend advanced systems intended for the software requirements description, analysis, and documentation with the capabilities of the Prolog logic programming language. The latter itself can be considered as fully universal and above all as an executable specification language. The background and some characteristic features of the SAM system developed within the frame of this approach are briefly described.

REFERENCES

[1] Boehm B.W. Software Engineering Economics. Englewood Cliffs, New Jersey: Prentice-Hall, 1981, 767 p.

[2] Teichroew D., Hershey E.A. PSL/PSA: A Computer-Aided Technique for Structured Documentation and Analysis of Information Processing Systems. IEEE Transactions on

Software Engineering. v.SE-3, 1, 1977, p.41-48.

[3] Сеничкин В.И. Языки спецификаций информационных систем. Измерения, Контроль, Автоматизация. 4(60), 1986, с.61-75.

[4] Кароль А.М. Система поддержки спецификации требований к программному обеспечению на базе логического программирования. Тез. докл. 2-ой международн. конференции "Программное обеспечение ЭВМ". Секция "Новая информационная технология". Калинин, 1987, с.73-76.

[5] Clocksin W. F., Mellish C. S. Programming in Prolog. Springer-Verlag, 2nd Edition, 1984.

[6] Gomaa H. The Impact of Rapid Prototyping on Specifying User Requirements. ACM SIGSOFT Software Engineering Notes. v.8, 2, April 1983, p.17-28.

[7] Knuth E., Rado P., Toth A. Preliminary Description of SDLA. MTA SZTAKI, Tanulmanyok, Budapest, 105/1980, 62 p.

[8] Knuth E., Habasz F., Rado P. SDLA - System Descriptor and Logical Analyzer. In: Information Systems Design Methodologies: A Comparative Review. T.W. Olle et al. (eds.), IFIP, North-Holland, 1982, p.143-171.

[9] Demetrovics J., Knuth E., Rado P. Computer-Aided Specification Techniques. Series in Computer Science, v.1, World Scientific, Singapore, 1985, 114 p.

[10] Toth A. A Software for Computer-Aided System Description. In: System Description Methodologies. D. Teichroew and G. David (eds.), IFIP, North-Holland, 1985, p.593-606.

[11] Borgida A., Greenspan S., Milopoulos J. Knowledge Reprezentation as the Basis for Requirement Specification. IEEE Computer, v.18, 4, 1985, p.82-90.

[12] Dubois E., Hagelstein J., Lahou E., Ponsaert F., Rifaut A. A Knowledge Reprezentation Language for Requirements Engineering. Proc. of the IEEE, v.74, 10, 1986, p.1431-1444.

[13] Demetrovics J., Knuth E., Rado P. Specification Meta Systems. IEEE Computer, v.15, 5, 1982, p.29-35.

[14] Yeh R.T., Zave P. Specifying Software Requirements. Proc. of the IEEE, v.68, 9, 1980, p.1077-1085.

[15] Brodie M.L. On the Development of Data Models. In: On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases and Programming Languages. Springer-Verlag, M.L. Brodie et al. (eds.), North-Holland, 1984, p.19-47.

[16] Rueher M. Quelques apports de PROLOG en Genie logiciel. BIGRE + Globule, 45, 1985, p.97-102.

[17] Campbell J.A. Knowledge-based Computing Systems and Software Engineering. Comput. Physics Commun., v.41, 2-3, 1986, p.285-290.

A KNOWLEDGE-BASED APPROACH TO SOFTWARE REQUIREMENTS SPECIFICATION

A.M. Karol

Summary

The paper describes the SAM system (Specification, Analysis, Modelling) which is a knowledge base that can be filled with knowledge and is intended for the development of software requirements and allows smooth transition to more detailed specifications. In many respects SAM continues the line "PSL/PSA-SDLA-SAD". For example, SAM does not fix previously defined technology of usage, and can serve as a tool for supporting different technologies. SAM is based on the ideas of conceptual database modelling and logic programming elements.

TUDÁSBÁZIS ALAPU KÖVETELMÉNY SPECIFIKÁCIÓ

A.M. Karol

Összefoglaló

A dolgozat az SAM /Specification, Analysis, Modelling/ rendszert tárgyalja. Az SAM egy szoftver követelmény specifikációt segitő tudásbázis, amely átmenetet biztosit a részletesebb specifikációhoz. Az SAM több vonatkozásban a "PSL/PSA-SDLA-SAD" vonal folytatásának tekinthető. Az SAM nem rögziti a használandó technológiát, többféle technológia támogatására képes. Az SAM fogalmi adatmodellezésen és a logikai programozás elvein alapszik.

MTA SZTAKI Közlemények 39/1988 p. 83-92.

SEPARATORY SUBSEMILATTICES AND THEIR PROPERTIES

L. O. LIBKIN

Computer Centre Moscow Institute of Steel and Alloys,
Leninskij Prosp., 4, 117936, USSR

I. B. MUCHNIK

Institute of Control Sciences, Profsoyuznaya,
65, Moscow, 117806, USSR

ABSTRACT

In this paper we study separatory subsemilattices of a semilattice, i.e. subsemilattices having set-teoretical complement as subsemilattice. The separation theorem is proved for arbitrary semilattice. We study the representation of separatory subsemilattices and the property of extension of the maximal chains in separatory subsemilattices. The real-valued function on semilattice connected to our main object and some generalizations are also discussed. The case of finite distributive semilattices is particularly considered.

1. INTRODUCTION

First we introduce the terminology. Let L (possibly with indices) be a semilattice, in other words, an algebra with one binary idempotent commutative associative operation denoted by point. A partial order on L is defined as $x \leq y$ iff $x \cdot y = y$ so L is join-semilattice. As usual, $J(L)$ denotes the set of all join-irreducible elements of L. An interval $[x,y]$ is the set $\{l \in L | x \leq l \text{ and } l \leq y\}$.

The semmilattice is said to be distributive (Grätzer [1978]) if for $x \leq y \cdot z$ there are $y' \leq y$ and $z' \leq z$ such that $x = y' \cdot z'$. It is an easy observation that finite distributive semilattice is distributive lattice.

The motivation of this paper as follows. The set of all subsemilattices of a semilattice is a convex geometry (Edelman, Jamison [1985], Edelman [1986]), which was introduced as an abstract model of convexity. One of the brances in the abstract convex analysis deals with different separation properties of the convex sets (Soltan [1984], Van de Vel [1984, axioms S_1-S_4]). The strongest separation property is following: the convex sets with empty intersection can be separated by halfspace. This property had been introduced and studied in Ellis [1952]. The halfspace in that paper had been defined as a convex set having a convex set as set-theoretical complement.

Now we can give our main definition.

Definition 1. The subsemilattice L' of the semilattice L is called separatory if $L-L'$ is also a subsemilattice of L.

We shall show that separation theorem holds for arbitrary semilattice. The other sections of the paper deal with different properties of separatory subsemilattices.

Notes. 1. There are other models of convexity having subsemilattices as examples. E.G. Romanowski and Smith [1985] introduced the concept of moda (entropic idempotent algebra). The semilattice is the simplest example of moda. The other example is the family of all convex subsets of Euclidean Space.

Korte and Lovász [1984] studied upper interval greedoids, or APS-greedoids. One can easily prove that the systems of complements of subsemilattices (in finite case) is APS-greedoid.

2. There is other reason for studyind the separatory subsemilattices, which is connected to applied tasks of data analysis (Libkin, Muchnik, Schwartzer, to appear).

3. Here we do not give proofs of theorems. They will appear elsewhere.

2. THE ALGEBRAIC CHARACTERIZATION OF SEPARATORY SUBSEMIALTTICES

For arbitrary semilattice L the set of all its subsemilattices partially ordered by set-theoretical inclusion is a lattice denoted Sub L.

Theorem 2.1. The subsemilattice L' of L is separatory iff L' has a pseudocomplement in lattice Sub L.

3. SEPARATION THEOREM

In this section L is arbitrary semilattice.

Theorem 3.1. (Separation Theorem). If L_1 and L_2 are subsemilattices of L with empty intersection then there is a separatory subsemilattice L' such that $L_1 \subseteq L'$ and $L_2 \subseteq L-L'$.

In other words, the subsemilattice L_1 can be separated from L_2 if and only if $L_1 \cap L_2 = \emptyset$.

The separation theorem has some corollaries.

Corollary 3.2. Any subsemilattice L_1 of semilattice L is the intersection of separatory subsemilattices of L.

Corollary 3.3. If subsemilattice $L' \subseteq L$ is meet-irreducible the lattice Sub$L$ then L' is separatory.

Note that Corollary 3.2 is one of the separation properties for abstract convexity: any convex set is an intersection of halfspaces (axiom S_3 in Van de Vel [1984]).

Corollary 3.4. Let L_1 be a subsemilattice of L. There is a separatory subsemilattice $L' \subseteq L$ such that $L_1 \cap L' = \{a\}$ iff $a \in J(L_1)$. Moreover, one can find L' with element a as unit.

4. FINITE DISTRIBUTIVE SEMILATTICES

In this case we find the construction of separatory subsemilattices and give the represenatation of subsemilattices according to corollary 3.2.

Theorem 4.1. 1). The subset L_o of finite distributive semilattice L is a subsemilattice if

$$L_o = L - U([x_i, y_i] \mid i \in I), \ \forall i \in I : x_i \in J(L). \qquad (1)$$

2). L_o is separatory subsemilattice of L iff it can be expressed in form (1) and $\{y_i \mid i \in I\}$ is a chain in L.

Consider $L_{x,y} = L - [x,y]$, $x \in J(L)$.

$L_{x,y}$ is the separatory subsemilattice and (1) shows that every subsemilattice L is an intersection of subsemilattices defined as $L_{x,y}$.

5. THE MAXIMAL CHAINS PRESERVATION PROPERTY

Fist we give the usual definition of a fragment in a chain. Let C be a chain in L and $C'\subseteq C$. C' is called the fragment of C if for minimal and maximal element c_0 and c_1 of C' the intersection $[c_0,c_1]\cap C$ is equal to C'.

In this section we consider only finite semilattices.

Definition 5.1. A finite semilattice L is called MCP -semilattice if for every separatory subsemilattice $L'\subseteq L$, maximal chain C' in L' and its extension to the maximal chain C in L the chain C' is a fragment of C. MCP is the abbreviation for Maximal Chains Preservation .

This definition has another interpretation. For any chain C in L one can define the binary relation R_C on C as follows: $(a,b)\in R_C$ iff $a>b$ and if $c\in C$, $a\geq c\geq b$ then $a=c$ or $b=c$. The definition 5.1. says that R_C , is a restriction of cover relation $\succ$ in L to C'.

As usual an open interval is an interval without maximal and minimal elements.

Theorem 5.2. The following statements are quivalent:

1). L is MCP-semilattice;

2). Any open interval is not a subsemilattice of L;

3). A three-elements chain cannot be an interval in L.

One can show that MCP as property of lattices is not connected to identities or quasi-identities.

Corollary 5.3. In every non-trivial variety of lattices both MCP-lattices and non-MCP-lattices exist.

For example, boolean, matroid and relative complemented lattices are MCP-lattices. The chains and finite free distributive lattices give examples.es of lattices without this property.

The separatory subsemilattices of MCP-semilattices possess a property like density or coherence .

Definition 5.4. The subsemilattice $L'\subseteq L$ is said to be dense if for every $x,y\in L'$ there are such $x = x_1,\ldots,x_n = y$ in L' that $x_i \sim x_{i+1}$; $i=1,\ldots,n-1$; here $x_i \sim x_{i+1}$ iff x_i covers x_{i+1} or x_{i+1} covers x_i.

Corollary 5.5. Every separatory subsemilattice of MCP-semilattice is dense.

Corollary 5.6. Every semilattice L can be embedded into a semilattice M such that image of L is an intersection of dense subsemilattices of M.

6. REPRESENTATION BY REAL-VALUED FUNCTION

According to Soltan [1984] the function $f:L\to\mathbb{R}^1$ is called quasiconvex if

$\forall x,y\in L : f(x\cdot y) \le max\{f(x),f(y)\}$,

and quasiconcave if

$\forall x,y\in L : f(x\cdot y) \ge min\{f(x),f(y)\}$.

The function both quasiconvex and quasiconcave is said to be quasilinear (Libkin, Muchnik, Schwartzer).

Consider the set Φ_L of all quasiconvex function on L. The Φ_L-convexity defined by Soltan [1984] is the set of all subsemilattices of L. In other words the following lemma holds.

Lemma 6.1. 1 . The level sets $L_\alpha=\{x\in L|f(x)\le\alpha\}$ of quasiconvex functions are subsemilattices of L.

2). The level sets $L^\alpha=\{x\in L|f(x)\ge\alpha\}$ of quasiconcave function are subsemilattices of L.

3). The level sets L_α and L^α of quasilinear function are separaotry subsemilattices of L.

The corollary 3.1. can be written in following form.

Theorem 6.2. Every bounded quasiconvex (or quasiconcave) function is supremum (accordingly an infimum) of the quasilinear functions.

The representation of functions on finite distributive lattices can be found.

Theorem 6.3. Consider a finite distributive semilattice L.

1). Function f is quasiconcave iff there is a function $\pi(a,x): J(L)\times L\to\mathbb{R}^1$ such that $a\le x\le y \Rightarrow \pi(a,x)\le\pi(a,y)$ and

$\forall x\in L : f(x)=min(\pi(a,x)|a\in J(x))$, (2)

where $J(x) = \{a\in J(L)|a\le x\}$.

This function π may be expressed as

$$\pi(a,x) = max(f(y)|y\in[a,x]). \quad (3)$$

2). Function f is quasilinear iff

$$\forall x \in L : f(x) = \min_{a \in J(x)} \max_{b \in J(x)} f(a \cdot b) = \max_{a \in J(x)} \min_{b \in J(x)} f(a \cdot b).$$

3). Conversely, if L is a lattice and f is quasiconcave iff (2) and (3) hold, then L is distributive.

In boolean case the results have simplest form.

Theorem 6.4. Consider $X = \{x_1, \ldots, x_n\}$. The setfunction $f: 2^x \to \mathbb{R}^1$ is quasilinear iff $\forall V \subseteq X : f(V) = \min_{x_i \in V} \max_{x_j \in V} a_{ij}$, and $A = \|a_{ij}\|$ is $n \times n$ -symmetrical matrix such that every submatrix of A has a saddle-point. Every quasiconcave function is a minimum of quasilinear ones.

Corollary 6.5. There is a one-to-one correspondence between separatory subsemilattices of 2^x and zero-one square matrices A which are symmetrical and such that every submatrix of A has a saddle-point.

Notes. 1). The theorem 6.4. shows that there is a connection between our main object and payoff matrices in game theory.

2). Corollary 6.5. gives a representation of separatory subsemilattices of sets by square zero-one matrices. An analogous representation of lattices, rings, topologies and algebras had been given in Ashman, Ficker [1984].

3). The theorem 6.4. and corollary 6.5. use the result proven by Gurvitch (Gurvitch, Libkin, to appear): if A is symmetrical matrix and every submatrix of A with equal sets of lines and columns has a saddle-point then every submatrix has a saddle-point (so called absolute definiteness of matrix).

7. SOME GENERALIZATIONS

For every signature one can give definition of separatory subalgebra. Consider an algebra (A, Ω), where A is a set and Ω is a signature.

Definition 7.1. The subalgebra (B, Ω) of (A, Ω) is said to be separatory if $(A-B, \Omega)$ is also subalgebra of (A, Ω).

For idempotent algebras one can give a characterization of separatory subalgebras like in semilattice case.

Theorem 7.2. The subalgebra (B, Ω) of an idempotent algebra (A, Ω) is separatory iff (B, Ω) has a pseudocomplement in lattice of subalgebras of (A, Ω).

This fact immediately implies that atomic algebraic pseudocomplemented lattice is boolean.

For signature Ω the separation theorem can be written as $ST(\Omega)$: for every algebra (A,Ω) and its subalgebras (B,Ω) and (C,Ω) with empty intersection there is a separatory subalgebra (S,Ω) such that $B\subseteq S$ and $C\subseteq A-S$.

In the section 3 we proved that $ST(\Omega)$ holds if Ω is the signature of semilattice. If Ω is the signature of group $ST(\Omega)$ does not hold, because there are not subgroups with empty intersection. It is easy to see that $ST(\Omega)$ does not hold also in semigroup case.

Let point denotes binary operation on set G, $X\subseteq G$ is a finite, subset, $\leq_X$ is a linear order on X and $x_1\leq_X\ldots\leq_X x_n$.
For $x\in A$ we define $p(x,X)$ as $(\ldots(x\cdot x_1)\cdot x_2\ldots)\cdot x_n$.

Theorem 7.3. The idempotent commutative groupoid G ensures separation theorem iff for every fintie $X,Y\subseteq G$ with empty intersection, $x\in G-(X\cup Y)$ and linear orders $\leq_X$ and $\leq_Y$ the subgroupoids generated by $X\cup\{p(x,Y)\}$ and $Y\cup\{p(x,X)\}$ have nonempty intersection.

Corollary 7.4. Idempotent commutative nonassociative groupoid ensuring separation theorem exists.

The last result shows that unlike semilattice case ST does not hold for lattices.

Theorem 7.5. If lattice has less than 5 elements then ST holds. For every $n\geq 5$ there is such lattice Z that ST does not hold and Z has n elements. For modular (distributive, boolean) lattices this limit is also 5 (accordingly 6 and 8).

REFERENCES

C.J. Ashman, V. Ficker: Families of sets on a finite set; J. Austral. Math. Soc., A-37 (1984) 405-412.

P.H. Edelman: Abstract convexity and meet-distributive lattices; Contemp. Math., 57 (1986) 127-150.

P.H. Edelman, R.E. Jamison: The theory of convex geometries; Geometriae Dedicata, 19 (1985) 247-270.

J.W. Ellis: A general set-separation theorem; Duke Math. J., 19 (1952) 417-421.

G. Grätzer: General lattice theory; Academie - Verlag, Berlin, 1978.

V.A. Gurovitch, L.O. Libkin: Quasi-linear set-functions and absolutely defined matrices (in Russian); Automatika i Telemehanika, to appear.

B. Korte, L. Lovász: Shelling structures, convexity and a happy end; Graph theory and Combinatorics: Proceeding of the Cambridge Combinatorical Conference in Honour of Paul Erdös (B. Bollobás, Ed.) Academic Press, New York/London, 1984, 219-232.

L.O. Libkin, I.B. Muchnik, L.V. Schwartzer: The quasi-linear monotone systems (in Russian); Automatika i Telemehanika, to appear.

I. Rivial: Maximal sublattices of finite distributive lattices. 2; Proc. Amer. Math. Soc., 44 (1974) 263-268.

A.B. Romanowska, J.D.H. Smith: Modal theory: An algebraic approach to order, geometry and convexity; Heldermann - Verlag, Berlin, 1985.

V.P. Soltan: Introduction to the axiomatic theory of convexity (in Russian); Stiinca, Kisinev, 1984.

M.Van de Vel: Binary convexities and distributive lattices; Proc. London Math. Soc., 48 (1984) 1-33.

SEPARATORY SUBSEMILATTICES AND THEIR PROPERTIES

L.O. Libkin, I.B. Muchnik

Summary

In the paper separatory subsemilattices of a semilattice are studied (a subsemilattice is separatory if its set-theoretical complement is also a subsemilattice). The separation theorem is proved for arbitrary semilattice. The representation of separatory subsemilattices and the property of extension of the maximal chains are also studied. Some real-valued functions connected to the main object are discussed. The case of finite distributive semilattice is particularly considered.

SZEPARÁLÓ RÉSZFÉLHÁLÓK ÉS TULAJDONSÁGAIK

L.O. Libkin, I.B. Muchnik

Összefoglaló

A félháló részfélhálója szeparáló, ha a halmaz elméleti komplementere is részfélháló.
A cikkben a szerzők egy szeparációs tételt bizonyitanak tetszőleges részhálóra. A cikk vizsgálja a szeparáló részfélhálók reprezentációját és a benne lévő maximális láncok kiterjeszthetőségét. A félhálókon definiált bizonyos valós értékü függvényekkel is foglalkozik a szerző, amelyeknek különleges szerepük van a tárgyaltakban. Különösen a véges disztributiv félhálókat tárgyalja részletesen.

MTA SZTAKI Közlemények, 39/1988 p. 93-100.

SEPARATION THEOREMS FOR LATTICES

Leonid O. LIBKIN

Computer Centre Moscow Institute of Steel and Alloys,
Leninskij Prosp. 4, Moscow 117936, USSR

ABSTRACT

In this paper the separation properties of sublattices of a lattice are studied. A sublattice of a lattice is said to be separatory if its set-theoretical complement is also a sublattice. The full characterization of lattices, satisfying the following property: two disjoint sublattices can be separated by separatory sublattice, is given. This property restricted to the intervals is equivalent to the distributivity of lattice. It is also shown that the sublattices of a lattice form a convex geometry (or their complements form an antimatroid) iff the lattice diagram does not contain a subdiagram like the letter N. The structure of separatory sublattices is destribed for finite distributive lattices.

1. INTRODUCTION

Let X be an arbitrary set and Z a family of subsets of X satisfying the properties:

O1. $\emptyset \in Z$, $X \in Z$.

O2. If $A \in Z$ and $B \in Z$ then $A \cap B \in Z$.

Z is called an alignment. The elements of Z are called closed or sometimes convex subsets of X. Given an alignment Z define the closure operator

$$Z(A) = \cap(B \mid A \subseteq B \in Z).$$

A subset A is closed iff $Z(A)=A$. The family of closed subsets, ordered by set-theoretical inclusion, is a lattice and $A \wedge B = A \cap B$, $A \vee B = Z(A \cup B)$.

Let $\langle L,\vee,\wedge \rangle$ be a lattice, $S(L)$ be a family of its sublattices and $Int(L)$ be a family of its intervals $[x,y] = \{l \in L \mid x \leq l \leq y\}$, $x \leq y$ iff $x \vee y = y$. Both $S(L)$ and $Int(L)$ are alignments.

According to abstract convexity theory (cf. Soltan [1984] Van de Vel [1984]) the convex subset is said to be halfspace if its set-theoretical complement is also convex.

Formerly we studied the halfspaces of semilattices. It was proved that two disjoin subsemilattices of a semilattice can be separated by halfspace, i.e. there is a halfspace containing the first subsemilattice and no element of the second one, see Libkin, Muchnik [1988]. According to the semilattice case, let us give the following definition.

<u>Definition 1.</u> A sublattice L' of a lattice L is called separatory if $L-L'$ is also sublattice of L.

The prime ideals give examples of separatory sublattices. The partition lattice Part (4) gives an example of lattice without separatory sublattices.

The separation theorem can be expressed as

ST. $\forall L_1, L_2 \in S(L): L_1 \cap L_2 = \emptyset \Rightarrow \exists L' \in S^*(L): \ L_1 \subseteq L', \ \ L_2 \subseteq L-L'$,

where $S^*(L)$ is the family of separatory sublattices.

This property ST had been introduced in the paper of Ellis [1955]. The elignment satisfying ST is said to be normal (cf. Soltan [1984]).

In the section 2 the characterization of lattices L having the normal alignment $S(L)$ is given.

In the section 3 the following modification of ST is studied:

STI. $\forall L_1, L_2 \in Int(L): L_1 \cap L_2 = \emptyset \Rightarrow \exists L' \in S^*(L): L_1 \subseteq L', \ L_2 \subseteq L-L'$.

It is proved that STI holds iff L is distributive. It generalizes the result of Van de Vel [1984].

The section 4 deals with the convex geometries. An aligmment Z on a finite set X is called convex geometry (Edelman, Jamison [1985], Edelman [1986]) if in addition C3 holds:

C3. If $A \in Z, x, y \notin A$ and $x \in Z(A \cup y)$ then $y \notin Z(A \cup x)$.

This axiom C3 is called an anti-exchange axiom, cf. Edelman [1980]. The convex geometries are equivalent to the important class of greedoids named the shelling structures, cf. Korte and Lovasz [1984]. In section 4 we describe the lattices such that $S(L)$ is a convex geometry.

In the last section we give the description of the separatory sublattices of finite distributive lattices.

According to Gratzer [1978] $J(L)$ and $M(L)$ stand for the sets of join-irreducible and meet-irreducible elements of L respectively.

C_2 denotes the two-element chain.

2. THE LATTICES SATISFYING SEPARATION THEOREM

It was shown in the paper of Libkin and Muchnik [1988] that for every $n \geq 5$ there is a lattice L such that $|L| = n$ and $S(L)$ is not normal, i.e. ST does not hold. For distributive lattices this limit is equal to 6. That statement allows to build only five lattices with normal alignment $S(L)$. Here we give a full characterization of these lattices.

Definition 2.1. Let L_1 be a lattice with unit e and L_2 be an arbitrary lattice. Then the ordinal sum of posets $L_1 - e$ and L_2 is called a weak ordinal *sum* of L_1 and L_2. We denote it by $L_1 + L_2$.

It is an easy observation that + is associative operation for bounded lattices. Therefore one can speak of a weak ordinal sum of the family of bounded lattices.

Theorem 2.2. Let L be an arbitrary lattice. $S(L)$ is a normal alignment (or ST holds) iff L is isomorphic to the weak ordinal sum of lattices each isomorphic to C_2 or $C_2^2 = C_2 \times C_2$.

Corollary 2.3. If ST holds for an arbitrary lattice L, then L is distributive and its width $w(L) \leq 2$.

3. THE LATTICES SATISFYING SEPARATION THEOREM FOR INTERVALS

It was proved that $Int(L)$ is a normal alignment iff L is distributive, see Van de Vel [1984]. In fact, the following holds.

Theorem 3.1. Given an arbitrary lattice L, the following are equivalent:

1). STI holds for L;

2). $Int(L)$ is a normal alignment;

3). L is distributive lattice.

The lattices of convex sets for $S(L)$ and $Int(L)$ are the lattice $SubL$ of sublattices of L and the lattice $CSubL$ of order--preserving sublattices of L repectively. It is known that in semilattice case the meet-irreducible element of lattice of the convex sets are separatory subsemilattices. Analogous statements are not true for $S(L)$ and $Int(L)$ in contrast to the semilattice case.

Theorem 3.2. If every sublattice from $M(SubL)$ is separatory, then L does not contain a sublattice isomorphic to the nondistributive five-element lattice M_5.

Theorem 3.3. Given a finite lattice L, the following are equivalent:

1). Every sublattice from $M(CSubL)$ is a separatory one;

2). $M(CSubL)$ contains exactly the prime ideals and filters of L;

3). L is distributive.

4. THE CONNECTION WITH CONVEX GEOMETRIES

Since the concept of convex geometry is used only in finite case (as well as the concept of antimatroid), in this section only the finite lattices are considered.

A finite lattice L is called N-free, if its diagram does not contain a subdiagram isomorphic to the diagram like the letter N or its dual, cf Rival [1986]. In other words, L is not N-free if there are four distinct elements $a,b,c,d \in L$ such that $a \wedge c = b$ and $b \vee d = c$..

Theorem 4.1. Given a finite lattice L, $S(L)$ is a convex geometry iff L is N-free.

Corollary 4.2. If alignment $S(L)$ is normal, then it is a convex geometry.

The connection between ST and the concept of convex geometry allows to give a characterization of a variety of destributive lattices.

Theorem 4.3. Let K be a nontrivial variety of lattices satisfying the following condition: if $S(L)$ is a convex geometry for the finite $L \in K$, $S(L)$ is normal. Then K is a variety of distributive lattices.

Theorem 4.4. Given a finite lattice L, an alignment $Int(L)$ is a convex geometry iff L is a chain.

5. SEPARATORY SUBLATTICE OF THE FINITE DISTRIBUTIVE LATTICES

Rival [1974] proved that the sublattices of a finite distributive lattice L and only they can be obtained as the subsets of L satisfying the following properties:

DL1. $L' = L - \cup([a_i, b_i] \mid i \in I)$,

DL2. $\forall i \in I$: $a_i \in J(L)$ and $b_i \in M(L)$.

Theorem 5.1. Let be a finite distributive lattice. $L' \subseteq L$ is a separatory sublattice of L iff it can be represented in form DL1, DL2 such that the following holds:

DL3. Both $\{a_i \mid i \in I\}$ and $\{b_i \mid i \in I\}$ are the chains.

One can also prove the inverse theorem.

Theorem 5.2. Let L be a finite lattice and DL1, DL2, DL3 describe exactly the separatory sublattices of L. Then L is distributive.

REFERENCES

P.H. Edelman: Meet-distributive lattices and the anti-exchange closure; Alg. Univ., 10 (1980) 290-299.

P.H. Edelman: Abstract convexity and meet-distributive lattices; Contemp. Math., 57 (1986) 127-150.

P.H. Edelman, R.E. Jamison: The theory of convex geometries; Geometriae Dedicata, 19 (1985) 247-270.

J.W. Ellis: A general set-separation theorem; Duke Math. J., 19 (1952) 417-421.

G.Gratzer: General Lattice Theory; Academie- Verlag, Berlin, 1978.

B. Korte, L. Lovász: Shelling structures, convexity and a happy end; Graph Theory and Combinatorics: Proceeding of the Cambridge Combinatiorcal Conference in Honour of Paul Erdős (B. Bollobás, Ed.), Academic Press, New York/London, 1984, 219-232.

L.O. Libkin, I.B. Muchnik: Separatory subsemilattices and their properties; MTA SZTAKI Közlemények, 39/1988.. p. 83-92.

I. Rival: Maximal sublattices of finite distributive lattices. 2; Proc. Amer. Math. Soc., 44 (1974) 263-268.

I. Rival: Stories about order and the letter *N*; Contemp. Math., 57 (1986) 263-285.

V.P. Soltan: Introduction to the axiomatic theory of convexity (Russian); Stiinca, Kisinev, 1984.

M. Van de Vel: Binary convexities and distributive lattices; Proc. London Math. Soc., 48 (1984), 1-33.

SEPARATION THEOREMS FOR LATTICES

L.O. Libkin

Summary

In the paper the separation properties of a sublattice of a given lattice are studied. A sublattice of a lattice is said to be separatory if its set-theoretical complement is also a sublattice. The full characterization of lattices satisfying the property: two disjoint sublattices can be separated by separatory sublattice, is given. It is also shown that the sublattices of a lattice form a convex geometry iff the lattice diagram does not contain a subdiagram like the letter N. The structure of separatory sublattices is described for finite distributive lattices.

HÁLÓKRA VONATKOZÓ SZEPARÁCIÓS TÉTELEK

L.O. Libkin

Összefoglaló

A cikkben az adott háló részhálóinak szeparációs tulajdonságairól van szó. Egy háló részhálója szeparáló, ha a halmaz-elméleti komplementere szintén részháló.
A szerző azon hálók teljes jellemzését adja meg, amelyek a következő tulajdonsággal rendelkeznek: két diszjunkt részháló szeparálható egy szeparáló részhálóval. Megmutatja, hogy egy háló részhálója akkor és csakis akkor konvex geometria, ha a háló-diagram nem tartalmaz egy N betüre hasonló részdiagramot. A szeparáló részhálók leirását is megadja, véges disztributiv hálók esetén.

MTA SZTAKI Közlemények 39/1988 p. 101-110.

ON A SUBSEMILATTICE-LATTICE OF A SEMILATTICE

L.O. LIBKIN

Computer Centre, Moscow Institute of Steel and Alloys
Leninskij Prosp. 4, Moscow, 117936, USSR

I.B. MUCHNIK

Institute of Control Sciences, Profsoyuznaya, 65,
Moscow, 117806, USSR

Abstract

In this paper a characterization of the lattices of all subsemilattices of semilattices is given. We also study the connections between the property of a semilattice and its subsemilattice-lattice. The structures of distributive, standard and neutral elements in this lattice are also found.

1. INTRODUCTION

The construction of a subsemilattice-lattice had appeared in our research on applied data analysis (Libkin, Muchnik, Schwartzer, to appear) and choice theory Libkin [1987]. In this paper we give a characterization of this object.

Consider a semilattice L (an algebra with one binary idempotent commutative associative operation). Let L be a join-semilattice with partial order defined as $x \leq y$ iff $x \cdot y = y$. (point denotes the operation of semilattice).

As usual $J(L)$ is the set of all join-irreducible elements of L.

Consider the set $S(L)$ of all subsemilattices of L. There are two operation $\wedge$ and $\vee$ on $S(L)$:

$L_1, L_2 \in S(L)$: $L_1 \wedge L_2 = L_1 \cap L_2$,

$L_1 \vee L_2 = L_1 \cup L_2 \cup \{l_1 \cdot l_2 \mid l_1 \in L_1, l_2 \in L_2\}$.

It is well-known that $\wedge$ and $\vee$ make $S(L)$ a lattice:

$$SubL = \langle S(L), \wedge, \vee \rangle.$$

The lattice thus constructed is called subsemilattice-lattice.

We shall also use the concept of pregeometry.

Definition 1.1. The pair $\langle A, C \rangle$ of a set A and closure operator C is called a pregeometry if all axioms of geometry hold, except the exchange axiom. Formally, C is an algebraic operator ($x \in C(X)$ iff for some finite $X_f \subseteq X$: $x \in C(X_f)$); $\forall x \in A$: $C(\{x\}) = \{x\}$ and $C(\emptyset) = \emptyset$.

Definition 1.2. For given closure operator C on A the operator $C_2 : 2^A \to 2^A$ is defined in the following way:

$$\forall X \subseteq A: \; C_2(X) = \cup (C(\{x,y\}) \mid x, y \in X).$$

2. THE CHARACTERIZATION OF SUBSEMILATTICE-LATTICE

Here we give two characterizations of a lattice $SubL$: as biatomic algebraic lattice and as the lattice of closed sets of a pregeometry.

Remind the definition. Atomic lattice $\mathcal{L}$ is said to be biatomic if, given an atom $p \le a \vee b$, there are atoms $a' \le a$, $b' \le b$ such that $p \le a' \vee b'$.

Consider the following properties of atomic lattice $\mathcal{L}$:

A1. Given the atoms $x_1, \ldots, x_n$ and atom $x \le x_1 \vee \ldots \vee x_n$, there are two atoms x', x'' lying in the principal ideals generated by no more than $n-1$ atoms from $\{x_1, \ldots, x_n\}$ such that $x \le x' \vee x''$.

A2. There is a semilattice operation on the set of atoms of $\mathcal{L}$ such that every principal ideal generated by two atoms is isomorphic to the subsemilattice-lattice of the semilattice generated by these two atoms.

Now we introduce two properties of a pregeometry $\langle A, C \rangle$.

B1. For every finite $X \subseteq A$ there is $n \in \mathbb{N}$ such that

$$C(X) = C_2^n(X).$$

B2. There is a semilattice operation on A such that $x \cdot y = z$ $(z \neq x,y)$ implies $C(\{x,y\}) = \{x,y,z\}$ and $x \cdot y \in \{x,y\}$ implies $C(\{x,y\}) = \{x,y\}$.

<u>Theorem 2.1.</u> The following statements are equivalent:

1) $\mathcal{L}$ is isomorphic to lattice $SubL$ for some L;
2) $\mathcal{L}$ is biatomic algebraic lattice and A1 holds;
3) $\mathcal{L}$ is atomic algebraic lattice and A1 and A2 hold;
4) $\mathcal{L}$ is isomorphic to the lattice of closed sets of a pre-geometry satisfying B1 and B2.

The scheme of the proof allows to obtain some corollaries about idempotent commutative groupoids.

<u>Corollary 2.2.</u> The lattice $\mathcal{L}$ is isomorphic to the lattice of subgroupoids of an idempotent commutative groupoid satisfying $(x \cdot y) \cdot y \in \{x,y,x \cdot y\}$ iff $\mathcal{L}$ is atomic algebraic lattice satisfying A1 and every principal ideal generated by two atoms contains no more than three atoms.

<u>Corollary 2.3.</u> The lattice $\mathcal{L}$ is isomorphic to the lattice of subgroupoids of an idempotent commutative groupoid satisfying $(x \cdot y) \cdot y = x \cdot y$ iff all conditions of the previous corollary hold and, in addition, for every four atoms a,b,c,d, $c \leq a \vee b$ implies that $c \vee d$ and b are not comparable.

We say that the class of algebras $\langle A,\Omega \rangle$ has the property (S) if lattice $\mathcal{L}$ is isomorphic to the lattice of subalgebras of $\langle A,\Omega \rangle$ iff it is algebraic lattice and every principal ideal of its compact element is isomorphic to the lattice of subalgebras of a finite algebra from this class. E.g. Grätzer, Koh and Makkai [1972] showed that the class of Boolean algebras has the property (S).

Corollary 2.4. The classes of groupoids described in two previous corollaries have the property (S).

Let Ω be a signature consisting only of idempotent operations, and (Ω) be the variety of algebras of this signature. Then atomic algebraic lattice $\mathcal{L}$ is isomorphic to the lattice of subalgebras of an algebra of (Ω) iff for every compact element $a \in \mathcal{L}$ the principal ideal $(a]$ is isomorphic to the lattice $Sub < A_a, \Omega>$, where A_a is the set of atoms less than a, and for two compact elements $a \leq b$ the following diagram is commutative

$$\begin{array}{ccc} (a] & \xrightarrow{\ \varphi\ } & (b] \\ {\scriptstyle i_1}\downarrow & & \downarrow{\scriptstyle i_2} \\ Sub<A_a,\Omega> & \xrightarrow{\ \Psi\ } & Sub<A_b,\Omega> \end{array}$$

where i_1 and i_2 are isomorphisms, φ and Ψ are natural embeddings.

Of course, semilattices satisfy this property, which is stronger than (S). But the use of A1 allows to make it weaker.

In fact, lattice $\mathcal{L}$ is isomorphic to $SubL$ iff A1 holds and the property, written above, holds for all elements a,b generated by no more than three atoms.

In the rest of this section we consider finite semilattices.

Note that the closure operator in semilattice case is antiexchange (Edelman [1980]). Therefore in the finite case the lattice $SubL$ is meet-distributive and its dual lattice is semimodular according to Edelman and Jamison [1985, p. 261]. The common properties of meet-distributive lattices can be found in Edelman [1986]. In particular, there is a height function on $SubL$: for $L_o \in SUBL$, $p(L_o)$ is the length of maximal chain from $\emptyset$ to L_o.

Corollary 2.5. 1) For every $L_o \in SubL$, $p(L_o)$ is equal to $|L_o|$. 2) For all atoms $x_1,\dots,x_n$ of $SubL$ $p(x_1 \vee \dots \vee x_n) \leq 2^n - 1$.

3) For every $L_1, L_2 \in SubL$:

$$p(L_1 \vee L_2) + p(L_1 \wedge L_2) \leq p(L_1) + p(L_2) +$$

$$+ (p(L_1) - p(L_1 \wedge L_2))(p(L_2) - p(L_1 \wedge L_2)).$$

3. FINITE DISTRIBUTIVE SEMILATTICES

Consider a finite semilattice L. L is called distributive if for any $x \leq y \cdot z$ there are $y' \leq y$, $z' \leq z$ such that $x = y' \cdot z'$ (for equivalent definitions and characterization see Rhodes [1975]).

The interval $[x,y]$ of L is said to be semisimple iff $x \in J(L)$.

Theorem 3.1. The subsemilattices of finite distributive semilattice L and only they can be written as $L - \cup([x_i, y_i] \mid i \in I)$, $\forall i \in I$: $[x_i, y_i]$ is a semisimple interval.

According to Rival [1974] consider the set of all semisimple intervals of L denoted by M. If $B \subseteq M$ then $F(B) =$ $= \{[a,l] \in M \mid [a,l] \subseteq U([x_i, y_i] \mid i \in I),\ \forall i \in I: [x_i, y_i] \in B\}$ and $F(L) = \{F(B) \mid B \subseteq M\}$

Corollary 3.2. F is a closure operator and $F(L)$ is a lattice. Moreover, $F(L)$ is dually isomorphic to $SubL$.

4. DISTRIBUTIVE, STANDARD AND NEUTRAL ELEMENTS IN $SubL$

The definitions of distributive, standard and neutral elements in lattice had been introduced by Ore, Grätzer and Birkhoff. One can find all definitions and main properties of these elements in Grätzer [1978]. In this section we give the structure of these elements in lattice $SubL$.

First we introduce two properties of subsemilattice $L_0 \subseteq L$.

$\forall x \notin L_0$: $L_0 \vee \{x\} = L_0 \cup \{x\}$. (1)

$\forall y \in L_0$: $(y = y_1 \cdot y_2$ and y_1, y_2 are noncomparable$) \Rightarrow y_1, y_2 \in L_0$. (2)

Theorem 4.1. The following statements are equivalent:

1) $L_0 \in SubL$ is distributive element;

2) $L_0 \in SubL$ is standard element;

3) (1) holds.

E.g. the filters are standard elements in $SubL$.

Theorem 4.2. $L_0 \in SubL$ is a neutral element if (1) and (2) hold.

Formerly (Libkin, Muchnik, to be published) we had introduced the concept of separatory subsemilattice. The subsemilattice $L_0 \leq L$ is said to be separatory if $L-L_0$ is also subsemilattice L.

Corollary 4.3. Every neutral element in lattice $SubL$ is separatory subsemilattice of L.

Corollary 4.4. Every neutral element of $SubL$ has a complement.

This corollary shows that there is a bijection between the neutral elements of $SubL$ and decompositions of $SubL$ into product of two lattices due to Grätzer [1978, Th. 3.4.1.]

5. THE CONNECTION BETWEEN PROPERTIES OF L AND $SubL$

In this section we find the semilattices L with boolean, pseudocomplemented, semimodular, planar and embeddable in free lattice subsemilattice-lattice $SubL$.

Theorem 5.1. The following statements are equivalent:

1) $SubL$ is a semimodular lattice;

2) $SubL$ is a pseudocomplemented lattice;

3) $SubL$ is a boolean lattice;

4) L is a chain.

This theorem shows that unlike lattice case(Lakser [1972]), only chains have a semimodular subsemilattice-lattice.

Theorem 5.2. The lattice $SubL$ can be embedded in a free lattice iff L is a chain and L has no more than three elements.

Theorem 5.3. The lattice $SubL$ is a planar lattice iff $|L| \leq 3$ and L is not a three-elements chain.

6. CONCLUSIVE REMARKS

Grätzer [1978] described the problem of characterization of sublattice-lattice. This problem was solved formerly in Rival [1974] for the finite distributive lattices in form like theorem 3.1 and corollary 3.2.

The equivalence of first, third and fourth statements in theorem 5.1 can be obtained as the consequence of the result of Shevrin [1963]. It was observed also in Salij, Kozyr [1972]. They also proved that algebraic atomic lattices are exactly the lattices of closed sets of pregeometrices.

Some other relationships between the properties of L and $SubL$ are the corollaries of theorem 4.1 and 4.2. E.g. these theorems and results of Grätzer, Schmidt [1958] entail that in finite case $SubL$ has a boolean congruence-lattice if and only if L is a chain.

The lattice $SubL$ is not unique algebra on set $S(L)$. The other algebra on $S(L)$ (bisemilattice) had been introduce and completely studied in Romanowska, Smith [1981].

REFERENCES

P.H.Edelman: Meet-distributive lattices and the antiexchange closure; Alg.Univ., 10 (1980) 290-299.

P.H.Edelman: Abstract convexity and meet-distributive lattices; Contemp. Math., 57 (1986) 127-150.

P.H.Edelman, R.E.Jamison: The theory of convex geometries; Geometriae Dedicata, 19 (1985) 247-270.

G.Grätzer: General lattice theory; Academie-Verlag, Berlin 1978

G.Grätzer, K.M.Koh, M.Makkai: On the lattice of subalgebras of a boolean algebra; Proc.Amer. Math.Soc.,36 (1972) 87-92.

G.Grätzer, E.T.Schimdt: Ideals and congruence relations in lattice; Acta Math. Acad.Sci.Hungar., 9 (1958) 137-175.

H.Lakser: A note on the lattice of sublattices of a finite lattice; Nanta Math., 6 (1973) 55-57.

L.O.Libkin: The algebraic methods for constructing and analysis of choice function classes (Russian); Trudy VNIISI, 14 (1978).

L.O.Libkin, I.B.Muchnik: Separatory subsemilattices and their properties; MTA SZTAKI Közlemények, 39/1988. p. 83-92.

L.O.Libkin, I.B.Muchnik, L.V.Schwartzer: Quasi-linear monotone systems (Russian); Automatika i Telemehanika, to appear.

J.B.Rhodes: Modular and distributive semilattices; Trans. Amer. Math.Soc., 201 (1975) 31-41.

I.Rival: Maximal sublattices of finite distributive lattices.2; Proc.Amer.Math.Soc., 44 (1974) 263-268.

A.B.Romanowska, J.D.H.Smith: Bisemilattices of subsemilattices; J.Algebra, 70 (1981) 78-88.

V.N.Salij, N.P.Kozyr: Geometric lattices and idempotent algebras (Russian);Uporiad. mnojestva i reshetki (Saratov) 2 (1974) 91-98.

L.N.Shevrin: The semigroups having modular subsemigroup-lattices (Russian); Dokl.Akad.Nauk USSR, 148 (1963) 292-295.

ON A SUBSEMILATTICE-LATTICE OF A SEMILATTICE

L.O. Libkin, I.B. Muchnik

Summary

In the paper a characterization of the lattices of all subsemilattices of semilattices is given.
The connections between the property of a semilattice and its subsemilattice-lattice are also studied.
The structures of distributive, standard and neutral elements in this lattice are also found.

EGY FÉL-HÁLÓ RÉSZFÉLHÁLÓJÁNAK HÁLÓJA

L.O. Libkin, I.B. Muchnik

Összefoglaló

A cikkben egy fél-háló részfélhálója hálójának a jellemzése van megadva. A szerző a félháló illetve az ő részfélhálójának hálója közötti összefüggést is tanulmányozza. Ezen háló strukturájáról, többek között a disztributiv, standard és neutral elemek strukturájáról is szó van.

MTA SZTAKI Közlemények 39/1988 p. 111-120.

SOFTWARE ENVIRONMENT FOR DISCRETE MATHEMATICAL RESEARCH IN UNIX-LIKE SYSTEMS

K.N. MANEV

Institute of Mathematics, Bulgarian Academy of Sciences

1. INTRODUCTION

The necessity of computers in the scientific researches is obvious nowadays. A lot of problems of the discrete mathematics, for example, need a computer treatment. The computer experiments became natural, preliminary or final, step of the investigation of some discrete configurations and thus, a part of the reasoning of new results. That is why the requirement towards software for mathematical researches are very strong.

In this paper a general aproach to software environment for discrete mathematical researces in UNIX-like systems is described. It sumarized, also, the autor's experience in implementation of some components of such an environment in UNIX System V and the obtained results.

2. UNIX-LIKE SYSTEMS

The choice of the opperating system is not accidental. The UNIX-like systems possess a lot of usefull features, some of them listed below:

- the UNIX-like systems are, practically, standartized nowadays and are installed in all wide-spread computer's architectures. This allows the exchange of software between researchers and minimizes the efforts for development;

- the UNIX-like systems have clear multvlevel structure (see Fig.) with appropriate interfaces in the differnt levels. so that any user (system analyst, researcher which **now** few programming or such that never will have been acquainted enough with the computer) can operate inside the environment on the corresponding level;

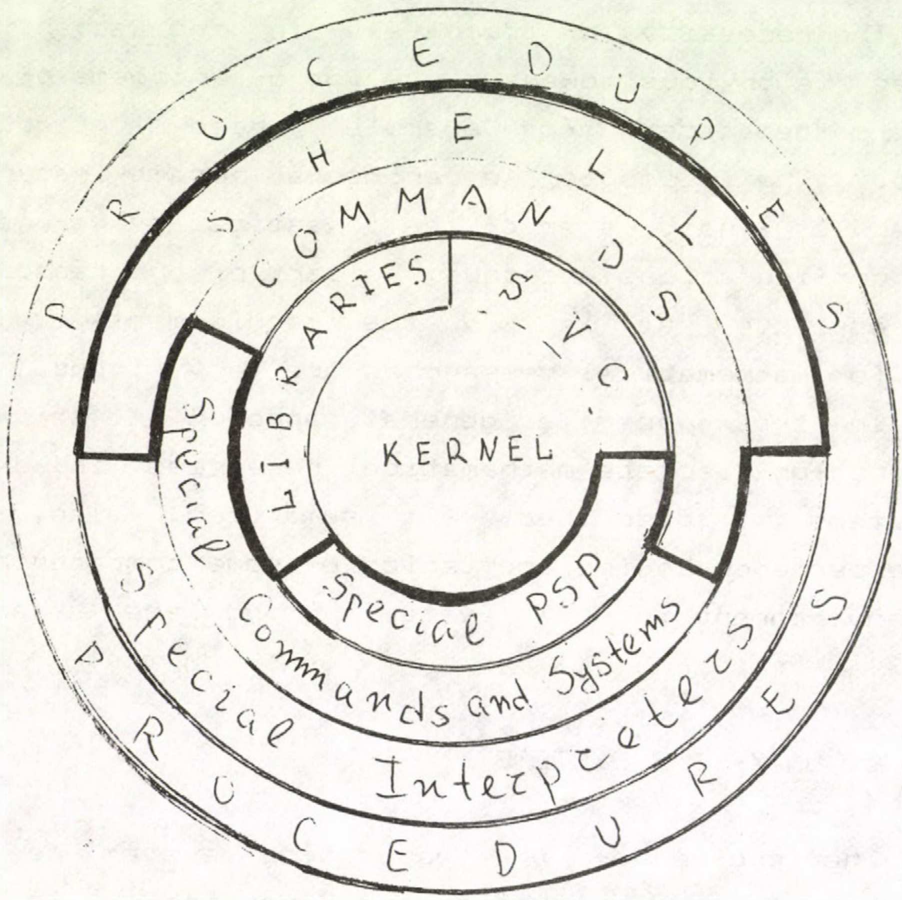


Fig.

- the UNIX-like systems support compilers, practically standartized, from the high-level programming language C, which allows effective usage of the computer's possibilities - pointers, bit and shift operations and so on. All dependences from the computer's architecture (the size of the computer's word, for example) can be defined as parametrs for the preprocessor of the language. So, the UNIX-like systems provide us with such portability of the software, which is of a big importance for our purposes:

- the UNIX-like systems communicate with users throw the command interpreter, which control running of the executable modules (commands). Any program, written from the, user becomes command. Moreover, the interpreter, itself, is also a command, and in a such a way is not obligatory. The user can provide an own interpreter. The standard command interpreter - shell, combines command with operators like this in C language: for, while, case and if. Whit the stancard commands echo (prints its arguments), read (reads values in its arguments), expr (evaluates an expression, formed by its arguments) and test (evaluates various predicates) shell becomes a programming language. Each procedure writen in this language is considered as a command, i.e. can be included in new procedures and so on.

3. SOFTWARE ENVIRONMENT

Our general idea is the next: to follow the multylevel structure of the operating system and to append on every level the corresponding tools for discrete mathematical researches (see Fig.)

The kernel of the UNIX-like systems is good enough for our purposes. It provides the supervisor calls (SVC), necessery for construction of an environment, dedicated to scientific researches of discrete mathematical objects. Moreover, appending of new SVC will decrease portability of the environment and the possibilities for exchanging of programs and results.

That is why we will not touch the kernel of the operating system.

The variety of architectures, operating systems and various versions, even for well standartized programming languages, did not permit wide-spread use of package standard programs (PSP). The UNIX-like systems give us the possibility for creating portable PSP. Moreover, the portability can be reached following a few simple rules:

- use only standard features of the C language;
- define any machine dependent constant on the preprocessor's level;
- use as subroutines only standard SVC, standard function from the system's libraries or programs from the package (which follow this rules, of course).

It is very helpfull to construct the package by the approach of abstract data type (ADT), i.e. to choose an ADT, appropriate for the goals of the package and to build the programs like operations with this ADT. In such a way, one can easy append new operations (programs) and change existing programs by ones, which implement better algorithms (and no needs to change the software which uses the package).

The PSP have to be created from very experienced system programers, which know very well the features of the operating system, C language and the object of the researches, too.

On the next level there are two possibilities. The first: any researcher, which knows the C language and can formulate a problem in the terms of the operations with the ADT, can write easy a program for resolving the problem, which use the PSP and with the help of a more experienced programer to create even complex systems, with approoriates languages, for resolving a class of problems.

The second possibility is: some standard programs to be transformed into commands of the operating system.

Although the user's systems and the commands lie in the same level of the structure of the operating system we will distinguish them. The commands are more simple and, as a rule, they have not a proper language. They perform usually one step of the resolving of the problem. The using of commands will be discussed bellow. The transformation of the standard programs into commands is more then trivial by the provided scheme for processing the arguments.

On the next level we propose the construction of new command interpreters. The standard command interpreter of the UNIX-like systems - shell, combines commands with the control constructions of the C language - for, while, case and if. For researchers, which dont know the C language there is only one possibility to work on their problems - sequential executing of commands. But these users know better others ways of composition of primitives - superposition of formulae, logical inference and so on.

That is why we propose on this level construction of new command interpreters, more appropriate for mathematicians. The UNIX-like systems encourage such activities, providing tools for automatically generating lexical analizers and parsers, starting from the corresponding formal gramars. Having such interpreters and enough commands (created on the previous level) for the special subject, even not experienced in programming user can formulate problems for resolving by computer.

What's about the mathematicians, which have not experience to work with computers at all. For these users we will create procedures in the languages of the command interpreters and the users could resolve their problems executing this procedures (giving appropriate values to the parameters).

Thus, we obtain an environment corresponding to the structure of the opperating system, which give to any user the possibility to determine the proper level of experience and to work on his problems with the appropriate tools. This environment is very flexible. It can be easy modified and developped, without changing the existing software.

4. SOME RESULTS

Some of our ideas were implemented in UNIX System V on the 32-bit (M68010) computer, with 2 Mb opperating memory - ICL CLAN. Here we briefly discuss the obtained results.

The portable package standard programes COMPACK [1] for combinatorial researches was created arround the ADT called a generated binary matrix (GBM). This ADT is appropriate for representing a very large set of combinatorial objects - graphs, matroids, block-designs, error-correcting codes, etc. The package contains programs, which operate with GBM, dosen't matter which object they represent, and an amount of programs appropriate for combinatorial researches. Many programs from COMPACK were transformed into commands of the opperating system.

The package GFPACK for operations with elements of Galois fields, polynomials and matrices over Galois fields is now under develeopment.

Using COMPACK, a special system - LINCOR, was constructed for computer researches of error-correcting codes [2]. For this purpose the package was extended with programs, appropriate for the special subject and we have enogh experience to extend the package in other directions, too.

The system LINCOR was designed to escape the everyday, routine work. Nevertheless, serious problems were resolved with it, connected with the covering radius of error correcting codes. It was refuted the conjecture that all binary linear codes, containing all-ones vector meet the lower bound for the covering radius [3] and the covering radius of 73 from 76 cyclic bynari codes with length 33, 35 and 39 were found [4].

Three languages for new command interpreters was proposed [5]. The first of them is extracted from Backus' FP-systems and combine comands in a functional style. The second is extension of the language Prolog, which treat both, commands and files, like predicates. The third is an original language, which combines commands by the usual scheme "backtrack" and it is oriented toward the searching of combinatorial configurations. All these languages are under experimental implementation in UNIX System V.

REFERENCES

1. K. N. Manev, A Portable Package of Standard Combinatorial Programs, Proceedings of the Scientific Seminar "Problems of the Portability of Programming Tools", Varna, 1987.
2. K. N. Manev, LINCOR - A System for Linear Codes Researces, in "Mathematics and Education in Mathematics", Proceedings of the Sixteen Spring Conference of the Union of Bulgarian Mathematicians, Sunny Beach, 1987.
3. S. M. Dodunekov, K. N. Manev, V. D. Tonchev, On the Covering Radius of the (14,6) Binary Codes, Containig The All-ones Vector, IEEE, Trans. on Inform. Theory, 1988 (to appear).
4. E. Velikova, K. Manev, The Covering Radius of Cyclic Codes of Length 33, 35 and 39, Annuaire de l'Université de Sofia, t. 81, l. 1 - Mathematique, 1987 (to appear).
5. K. N. Manev, Automatization of the Combinatorial Researches, Candidate disertation, University of Sofia, 1987.

SOFTWARE ENVIRONMENT FOR DISCRETE MATHEMATICAL RESEARCH IN UNIX-LIKE SYSTEMS

K.N. Manev

Summary

In the paper a general approach to software environment for research in discrete mathematics in UNIX-like systems is described.

The paper summarizes the author's experience in implementation of some components of such an environment in UNIX System V and a review of results is presented. The programs written are in a great extent independent of the combinatorial structure of the problems. However better results can be achieved using some additional sub-bases for particular problems (say coding theory).

DISZKRÉT MATEMATIKAI KUTATÁSOK SZOFTVER KÖRNYEZETE UNIX TIPUSU RENDSZEREK HASZNÁLATÁVAL

K.N. Manev

Összefoglaló

A cikkben a szerző egy általános módszertant mutat be, amely diszkrét matematikai kutatások szoftver környezetének megteremtésére vonatkozik UNIX-tipusu rendszereket felhasználva. A szerzőnek gazdag tapasztalata van az UNIX System V rendszerek esetében.
A cikkben bizonyos eredmények áttekintése is megtalálható. A megirt programok nagymértékben függetlenek a konkrét kombinatorikai strukturától, de jobb szoftver környezet is kialakitható konkrét esetekben /pl. a kódelméletben/.

MTA SZTAKI Közlemények 39/1988 p. 121-130.

STRONGLY DOMINATING SETS OF VARIABLES

Iv. MIRTCHEV and Sl. SHTRAKOV

Department of Mathematics
Pedagogical Institute
Blagoevgrad, Bulgaria

In this paper we introduce and investigate strongly dominating and regular dominating sets of variables for the functions.

DEFINITION 1.[2]. A function $f(x_1,\dots,x_n)$ is said to depend on the variable x_i, $1 \le i \le n$ if there exist $n-1$ constants $c_1,\dots,c_{i-1},c_{i+1},\dots,c_n$ for the variables $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n$ such that the function $f(x_1=c_1,\dots,x_{i-1}=c_{i-1},x_{i+1}=c_{i+1},\dots,x_n=c_n)$ assumes at least two different values. R_f denotes the set of all the variables on which f depends.

DEFINITION 2.[2]. A set M, $M \subseteq R_f$ is called separable for f with respect to the set $N=\{x_1,\dots,x_l\} \subseteq R_f$ if there exist l constants $c_1,\dots,c_l$ such that $M \subseteq R_{f(x_1=c_1,\dots,x_l=c_l)}$. $S_{f,N}$ denotes the set of all the separable sets for f with respect to the set N.

When M is separable for f with respect to $R_f \setminus M$ then M is called separable for f. S_f denotes the set of all the separable sets for f.

DEFINITION 3.[1]. A set M, $M \subseteq R_f$ is called c-separable for f with respect to the set $N=\{x_1,\dots,x_l\} \subseteq R_f$ iff for every l constants $c_1,\dots,c_l$, $M \subseteq R_{f(x_1=c_1,\dots,x_l=c_l)}$ holds true. $S^{*}_{f,N}$ denotes the set of all the c-separable sets for f with respect to the set N.

When M is c-separable for f with respect to $R_f \setminus M$ then M is called c-separable for f.

DEFINITION 4.[1]. A set $M=\{x_1,\dots,x_m\} \subseteq R_f$ is called dominating of N, $N \subseteq R_f$, $(M \xrightarrow{d} N)$ if there exist m constants $c_1,\dots,c_m$ such that $(*)$ $N \cap R_{f(x_1=c_1,\dots,x_m=c_m)} = \emptyset$ and M is minimal with respect to this property. When M satisfies $(*)$ then it is called α-dominating of N, $(M \xrightarrow{\alpha d} N)$.

$M \xrightarrow{\overline{d}} N$ denotes that M is not dominating of N for f and

$M \overset{\overline{\alpha d}}{\longrightarrow} N$ denotes that M is not α-dominating of N for f.

DEFINITION 5.[3]. A set $P = \{x_1,\dots,x_p\} \subseteq R_f$ is called strongly dominating of Q , $Q \subseteq R_f$,($P \overset{sd}{\longrightarrow} Q$) if for f there exists an p-tuple $P^* = (c_1^*,\dots,c_p^*)$ such that P is dominating of Q with P^* and

$$\bigcup_{i=1}^{p} C_{x_i,Q} = Q \text{ where } C_{x_i,Q} = Q \setminus R_{f(x_i=c_i^*)} .$$

The set $C_{x_i,Q}$ is called active zone of x_i in Q. $P \overset{\overline{sd}}{\longrightarrow} Q$ denotes that P is not strongly dominating of Q for f.

When $P \overset{sd}{\longrightarrow} Q$ and for any i,j , $1 \le i,j \le p$

$$C_{x_i,Q} \cap C_{x_j,Q} = \emptyset$$

the set P is called regular dominating of Q for f ($P \overset{rd}{\longrightarrow} Q$).

$P \overset{\overline{rd}}{\longrightarrow} Q$ denotes that P is not regular dominating of Q for f.

We now present an example to illustrate these definitions.

EXAMPLE 1. Let $f = x_1x_5 + x_2x_5 + x_3x_5 + x_3x_5\bar{x}_6 + x_4x_6 \pmod 2$, $P = \{x_5,x_6\}$ and $Q = \{x_1,x_2,x_3,x_4\}$. It is obvious that $P \overset{sd}{\longrightarrow} Q$ with $P^* = (0,0)$ and $C_{x_5,Q} = \{x_1,x_2,x_3\}$, $C_{x_6,Q} = \{x_4\}$. Since $C_{x_5,Q} \cap C_{x_6,Q} = \emptyset$ it follows that $P \overset{rd}{\longrightarrow} Q$.

THEOREM 1. If $P \overset{sd}{\longrightarrow} Q$ then for every $x_i \in P$, $C_{x_i,Q} \not\subseteq M$, where $M = \bigcup_{\substack{x_j \in P \\ x_j \ne x_i}} C_{x_j,Q}$.

PROOF. Obviously $P \overset{d}{\longrightarrow} Q$. Without loss of generality assume that $P = \{x_1,\dots,x_p\}$, $Q = \{x_{p+1},\dots,x_q\}$, $p < q \le n$. Now suppose that the theorem is false and let for example

$$C_{x_1,Q} \subseteq C_{x_2,Q} \cup \dots \cup C_{x_s,Q} , \; s \le p .$$

Consequently there exist s-1 constsnts $c_2^*,\dots,c_s^*$ such that

$$C_{x_1,Q} \cap R_{f(x_2=c_2^*,\dots,x_s=c_s^*)} = \emptyset.$$

This implies

$$Q \cap R_{f(x_2=c_2^*,\dots,x_p=c_p^*)} = \emptyset$$

for every p-s constants $c_{s+1}^*,\dots,c_p^*$ for the variables from the set.

$P \setminus (Q \cup \{x_1\})$. So we have $P\setminus\{x_1\} \xrightarrow{\alpha d} Q$ and $P \xrightarrow{\bar{d}} Q$. A contradiction. The theorem is proved.

COROLLARY 1. If $P \xrightarrow{sd} Q$ then for every $x_j, x_i \in P$, $x_i \neq x_j$

$$C_{x_i,Q} \not\subset C_{x_j,Q}.$$

COROLLARY 2. If $P \xrightarrow{sd} Q$ then for every $x_i \in P$, $C_{x_i,Q} \neq \emptyset$.

THEOREM 2. If $P \xrightarrow{sd} Q$ then $Card(P) \leq Card(Q)$.

This theorem follows by Theorem 1.

$P \xleftrightarrow{sd} Q$ denotes that $P \xrightarrow{sd} Q$ and $Q \xrightarrow{sd} P$.

PROPOSITION 3. If $P \xleftrightarrow{sd} Q$ then $Card(P) = Card(Q)$.

THEOREM 4. If $P \xrightarrow{sd} Q$ and $Card(P) = Card(Q)$ then $P \xrightarrow{rd} Q$ and $Q \xrightarrow{d} Q$.

PROOF . By $P \xrightarrow{sd} Q$, $Card(P) = Card(Q)$ and Theorem 1 it follows that there exist $Card(P)$ different active zones in Q which contain at least one variable belonging to only one active zone. Hence for every $x_i \in P$ it is true $Card(C_{x_i,Q}) = 1$. This implies $P \xrightarrow{rd} Q$.

Now suppose that $Q \xrightarrow{\bar{d}} Q$. Obviously $Q \xrightarrow{\alpha d} Q$ and there exists a subset Q_1 of Q for which $Q_1 \xrightarrow{d} Q$. Without loss of generality assume that $P = \{x_1, \ldots, x_p\}$, $Q = \{x_{p+1}, \ldots, x_{2p}\}$, $2p \leq n$ and $C_{x_i,Q} = \{x_{p+i}\}$ for $i=1,2,\ldots,p$. Let $x_{p+t} \in Q \setminus Q_1$ for some $t \leq p$. Consequently $C_{x_t,Q} \cap Q_1 = \emptyset$. As in Theorem 1 it follows that

$P\setminus\{x_t\} \xrightarrow{\alpha d} Q$ and $P \xrightarrow{\bar{d}} Q$. This contradiction completes the proof of the theorem.

COROLLARY. If $P \xrightarrow{sd} P$ then $P \xrightarrow{rd} P$.

In [1] the s-system of a family Ω of sets is introduced as follows: A set $\Sigma = \{x_1, \ldots, x_t\}$ is called s-system of $\Omega = \{P_1, \ldots, P_n\}$, $P_i \neq \emptyset$ if $\Sigma \cap P_i \neq \emptyset$, $i=1,2,\ldots,n$ and for every $j \leq t$ there exists $P_k \in \Omega$ such that $P_k \cap \Sigma = \{x_j\}$.

THEOREM 5. Let $P \xrightarrow{sd} Q$. The following two conditions are equivalent:

(i) $Card(P) = Card(Q)$

(ii) Q is an s-system of the family $\{C_{x_i,Q} | x_i \in P\}$.

PROOF. (ii)==>(i). Sinse Q is an s-system of $\{C_{x_i,Q} | x_i \in P\}$ it follows that for every $x_j \in Q$ there exists $x_i \in P$ such that $C_{x_i,Q} \cap Q = \{x_j\}$. By the definition of $C_{x_i,Q}$ it follows that $C_{x_i,Q} \subseteq Q$. Consequently $C_{x_i,Q} = \{x_j\}$. So every $x_j \in Q$ is an active zone of some $x_i \in P$. By Corollary 1 of Theorem 1 it follows that

$$Card(P) = Card(Q).$$

(i)==>(ii). Sinse $Card(P) = Card(Q)$ then as in the proof of Theorem 4 it follows that $Card(C_{x_i,Q}) = 1$ for every $x_i \in P$. So, any $x_j \in Q$ is an active zone of some $x_i \in P$. Hence Q is an s-system of the family $\{C_{x_i,Q} | x_i \in P\}$.

It is easy to prove that Q is unique s-system of the family $\{C_{x_i,Q} | x_i \in P\}$.

COROLLARY. If $P \xrightarrow{sd} Q$ and Q is an s-system of the family $\{C_{x_i,Q} | x_i \in P\}$ then $Q \xrightarrow{d} Q$.

EXAMPLE 2. Let $f = x_1x_4 + x_2x_3\bar{x}_4x_5 \pmod 2$, $P = \{x_4, x_5\}$, $Q = \{x_1, x_2, x_3\}$ and $P^* = (0,0)$. It is easy to see that $P \xrightarrow{sd} Q$, $P \xrightarrow{rd} Q$ but $P \xrightarrow{\bar{d}} P$ and $Q \xrightarrow{\bar{d}} Q$.

THEOREM 6. If $P = \{x_1, \ldots, x_p\} \subseteq R_f$ and $P \xrightarrow{d} P$ then $P \xrightarrow{rd} P$.

PROOF. Let x_i be an arbitrary variable from P. Obviously $\{x_i\}$ is dominating of $\{x_i\}$ for any constant c_i. Hence $x_i \in C_{x_i,P}$, $C_{x_i,P} \neq \emptyset$ and $P \subseteq \bigcup_{x_i \in P} C_{x_i,P}$.

On the other hand $\bigcup_{x_i \in P} C_{x_i,P} \subseteq P$. So, we obtain $P \xrightarrow{sd} P$ with any $P^* = (c_1^*, \ldots, c_p^*)$ and by Corollary of Theorem 4 it follows that $P \xrightarrow{rd} P$.

COROLLARY. If $P \xrightarrow{d} P$ then for each subset P_1 of P it is true $P_1 \xrightarrow{rd} P_1$.

THEOREM 7. If $P \xrightarrow{sd} Q$ and $P \cap Q = \emptyset$ then each non-empty subset P_1

of P is separable for f with respect to $P\setminus P_1$.

PROOF. Let $P=(x_1,\dots,x_p)$, $P_1=(x_1,\dots,x_s)$, $s\le p$ and $P^*=(c_1^*,\dots,c_p^*)$ be an p-tuple with which P is strongly dominating of Q for f.

Suppose the theorem is false i.e. $P_1\notin S_{f,P\setminus P_1}$. Then for any p-s tuple $(c_{s+1},\dots,c_p)$ it is true

$$P_1\setminus R_{f(x_{s+1}=c_{s+1},\dots,x_p=c_p)}\neq\emptyset .$$

Hence

$$P_2=P_1\setminus R_{f(x_{s+1}=c^*_{s+1},\dots,x_p=c^*_p)}\neq\emptyset$$

Let $x_i\in P_2$. We have

$$x_i\notin R_{f(x_{s+1}=c^*_{s+1},\dots,x_p=c^*_p)}$$

and

$$C_{x_i,Q}\cap R_{f(x_{s+1}=c^*_{s+1},\dots,x_p=c^*_p)}=\emptyset.$$

This implies

$$C_{x_i,Q}\subseteq\bigcup_{k=s+1}^{p}C_{x_k,Q}.$$

This contradicts to Theorem 1.Consequently $P_1\in S_{f,P\setminus P_1}$.The theorem is proved.

COROLLARY 1. If $P\xrightarrow{sd}Q$ and $P\cap Q=\emptyset$ then for any subset P_1 of P the set $(Q\setminus(\bigcup_{x_k\in P\setminus P_1}C_{x_k,Q}))\cup P_1$ is separable for f with respect to $P\setminus P_1$.

COROLLARY 2. If $P\xrightarrow{sd}Q$, $P_1,P_2\subseteq P$ and $P_1\cap P_2=\emptyset$ then $P_1\in S_{f,P_2}$.

THEOREM 8. Let $P=(x_1,\dots,x_p)\subseteq R_f$ and $P^*=(c_1^*,\dots,c_p^*)$ .If $Q=R_f\setminus(T\cup P)$, where $T=R_{f_1}$ and

$$f_1=f(x_1=c_1^*,\dots,x_p=c_p^*)$$

then $T\in S^*_{f,Q}$.

PROOF. Suppose the theorem is false. Then there exist q constants $c_{i_1},\dots,c_{i_q}$ for the variables in Q such that $T\not\subset R_{f_2}$,where

$$f_2=f(x_{i_1}=c_{i_1},\dots,x_{i_q}=c_{i_q}).$$

Without loss of generality assume that $x_r \in T\setminus R_{f_2}$. Obviously $Q \cap R_{f_1}=\emptyset$. Consequently for every q constants $d_{i_1},\dots,d_{i_q}$ it is true

$$f_1=f_1(x_{i_1}=d_{i_1},\dots,x_{i_q}=d_{i_q}).$$

This implies that

$$f_1=f_1(x_{i_1}=c_{i_1},\dots,x_{i_q}=c_{i_q}).$$

Now $x_r \notin R_{f_2}$ shows that $x_r \notin R_{f_1}$. Remember that $T = R_{f_1}$. Thus we have $x_r \notin T$. A contradiction. The theorem is proved.

COROLLARY 1. If $P \xrightarrow{sd} Q$ then $Q\setminus C_{x_i,Q} \in S^{*}_{f,C_{x_i,Q}}$ and $C_{x_i,Q} \in S^{*}_{f,M}$, where $M = Q\setminus C_{x_i,Q}$ for any $x_i \in P$.

COROLLARY 2. If $P \xrightarrow{rd} Q$ then for every $x_i, x_j \in P$, $i \neq j$ it is true $C_{x_j,Q} \in S^{*}_{f,C_{x_i,Q}}$.

THEOREM 9. If $P \xrightarrow{sd} Q$ then $P \xrightarrow{sd} Q \cup P$.

PROOF. Obviously $P \xrightarrow{d} Q \cup P$. Then for each variable $x_i \in P$ there exists an active zone of x_i in Q such that $C_{x_i,Q} \subseteq Q$. Thus we have $\{x_i\} \xrightarrow{d} C_{x_i,Q}$. By Corollary 2 of Theorem 7 it follows that $C_{x_i,M}=\{x_i\} \cup C_{x_i,Q}$ and $\bigcup_{x_i \in P} C_{x_i,M} = M$, where $M = P \cup Q$. The theorem is proved.

THEOREM 10. If $P \xrightarrow{sd} Q$ and $P \cap Q =\emptyset$ then for every $x_i \in P$ and for every $x_j \in C_{x_i,Q}$ there exists at least one m-set M, $Card(M)=m$ which is separable for f and $x_i, x_j \in M$.

This theorem follows by Theorem 4.10 [1] and Theorem 6.1 [6].

COROLLARY 1. If $P \xrightarrow{sd} Q$ then each variable $x_i \in P$ forms a separable pair with each variable $x_j \in C_{x_i,Q}$.

Denote by $\delta(x_i, L)$ the number of all the separable pairs which are formed between x_i and the variables from the set $L \subseteq R_f$ and by

$\delta(P,Q)$ the number of all the separable pairs from P X Q.

COROLLARY 2. If $P \xrightarrow{sd} Q$ then for any $x_j \in Q$

$$\delta(x_j, P) \geq Card(\{C_{x_i,Q} | x_j \in C_{x_i,Q} \& x_i \in P\}).$$

COROLLARY 3. If $P \xrightarrow{sd} Q$ then

$$\delta(P,Q) \geq \sum_{x_i \in P} Card(C_{x_i,Q}).$$

COROLLARY 4. If $P \xrightarrow{rd} Q$ then

$$\delta(P,Q) \geq Card(Q).$$

R E F E R E N C E S

1. Shtrakov Sl. Vl., Dominating and Annuling Sets of Variables for the Functions. Blagoevgrad, 1987.

2. Čimev K.N., Separable Sets of Arguments of Functions. MTA SzTAKI Tanulmanyok 180/1986.

3. Mirtchev I.A., Separable and Dominating Sets of Variables for the Functions. in Proc. of East European Category Seminar, Predela, 29.02-04.03. 1988.

4. Деметрович Я., Дьепеши Д., Генерирование функциональных зависимостей и их представление помощью реляции. MTA SzTAKI Kozlemenyek, 24, 1980, 7-18.

5. Деметрович Я., З. Фюреди и Д. Катона, Зависимости в составных базах данных. Кибернетика, No 5, 1985, 107-111.

6. Чимев К. Н., Отделими множества от аргументи на функциите. Благоевград, 1982.

7. Чимев К. Н., Функции и графи. Благоевград, 1983.

8. Мирчев И. А., Доминиращи, самодоминиращи и отделими множества от променливи на функциите. Сб. "Дискр. матем. и приложения", Благоевград, 1987.

STRONGLY DOMINATING SETS OF VARIABLES

Iv. Mirtchev, Sl. Shtrakov

Summary

The paper studies strongly dominating and regular dominating sets of variables of functions. It is a continuation of the research of the authors and K. Chimev. If is proved that if P strongly dominates Q for f, then all subsets $P_1 \subset P$ are separable for f with respect to $P \setminus P_1$.

A VÁLTOZÓK ERŐSEN DOMINÁLÓ HALMAZAI

Iv. Mirtchev, Sl. Shtrakov

Összefoglaló

A cikk a függvények változóinak un. erősen domináló és reguláris an domináló halmazaival foglalkozik /a pontos definiciók a cikkben megtalálhatók/. Ez folytatása a szerzők illetve K. Csimev korábbi kutatásainak. Be van bizonyitva, hogy ha P erősen dominálja Q-t /egy f függvényre nézve/ akkor minden $P_1 \subset P$ sz f-nek a $P \setminus P_1$-re nézve vett szeparábilis részhalmaza.

MTA SZTAKI Közlemények 39/1988 p. 131-150.

TWO TYPES OF RANDOM POLYHEDRAL SETS*

B. UHRIN

Computer and Automation Institute,
Hungarian Academy of Sciences
1502 Budapest, Pf. 63., Hungary

Let $A = \{a_{ij}\}_{i=1,j=1}^{m\ n}$, $\alpha = (\alpha_1, \ldots, \alpha_m)$, be a matrix and a vector (with real coefficients). By a (closed) polyhedral set (PS) we mean the solution set of the system of linear inequalities

(1) $\qquad Ax \geqslant \alpha$

i.e.

(2) $\qquad PS = \{x \in R^n : Ax \geqslant \alpha\}$

If (1) has no solution then we get the "empty polyhedral set", i.e. the empty set is included into the family of PS. Denote by $\mathcal{P}_m$ this family (when m and n are fixed and a_{ij} and α_i run through the real numbers). So the elements of $\mathcal{P}_m$ are all PS of the for m (2) plus the empty set.

One speaks of random polyhedral set (RPS), when a_{ij} and α_i are not real numbers but "real-valued" random variables. We shall assume that they are defined on a common probability-measure space (Ω, Σ, P), i.e. a_{ij} and α_i are real-valued measurable functions defined on Ω.

---

* Based on the author's dissertation: "Systems of linear inequalities, random polyhedral sets and quasi-concave functions", Dissertation, Hungarian Academy of Sciences, Budapest, 1978 (in Hungarian).

The first problem concerning RPS is how to define them?

The study of "random" geometrical objects (points, lines, hyperplanes, circular segments, e.t.c.) has a long history (see M. Kendall-Moran [1], Moran [2],[3]).

On the other hand, in the last few decades the most general random sets have been investigated (Matheron [4], D. Kendall [5], Santaló [6]).

As to RPS, sometimes they are defined as PS where the constants a_{ij} and α_i are randomly generated real numbers (Schmidt-Mattheiss [7]). Another approach is to consider RPS as an intersection of a finite number of random (closed) halfspaces (Rényi-Sulanke [8], Schmidt [9]). A "dual" approach to this is when (a bounded) RPS ("random polytope") is defined as the convex hull of a finite number of points given randomly in R^n.

A special case of Rényi-Sulanke approach is when the (normal vectors of) halfspaces are deterministic and only their translations are random, i.e. a_{ij} are real numbers and α_i are random variables.

In all above concepts of RPS both m an n are arbitrary (but fixed) integers. If we relax this condition, if say, n is fixed but m is also a random integer (i.e. discrete random variable), then we come to a much more complicated concept of RPS. Examples of these are so called Poisson polyhedra (see, e.g., [4] pp. 155-185).

Below we hall restrict ourselves to the study of two possible types of RPS.

The first type is an analogy of the random number: while the latter is a real-valued measurable function on a probability space, the RPS is a "PS-valued" measurable function.

In the second approach a RPS is the collection of those $x \in R^n$ that solve the system (1) with a probability at least p (where $0 \le p \le 1$ is a fixed number). The first type RPS stem for the general concept of a random set.

The second one was initiated by the people working in stochastic programming (more exactly "chance constrained" stochastic programming, see Kall [10] for a detailed survey).

There is no similarity between these two types. To study them, completely different techniques are needed and even the questions arising in conection with them are completely different.
We have concentrated to basic questions concerning these two types: in the first type the question of measurability of a "PS-valued" function and in the second one the question of convexity of RPS.

1. RPS of the first type

Let $a_{ij}, \alpha_i, i=1,\dots,m, j=1,\dots,n,$ be Borel-measurable real functions on (Ω, Σ, P) and denote $A(\omega) = \{a_{ij}(\omega)\}_{i=1, j=1}^{m \; n}$, $\alpha(\omega) = (\alpha_1(\omega), \dots, \alpha_m(\omega))$, $\omega \in \Omega$.
Further denote

(3) $$X(\omega) := \{x \in R^n : A(\omega)\, x \geq \alpha(\omega)\}, \; \omega \in \Omega .$$

After defining in $\mathcal{P}_m$ a special topology we endowe $\mathcal{P}_m$

with a Borel σ-algebra and prove that $\mathcal{X}$ is a measurable function w.r.t. this σ-algebra. This will mean that $\mathcal{X}$ is a RPS.

Denote by $\mathcal{K}$ the family of all polytopes from R^n (the empty set is included in $\mathcal{K}$). We recall that a polytope is a bounded PS, so $\mathcal{K}$ is a collection of all bounded PS, where m is arbitrary. Denote by $\mathcal{G}$ the family of all open PS (where m is arbitrary, empty set is included in $\mathcal{G}$). (An open PS is the set of the form $\{x \in R^n: \; Ax > \alpha\}$, m is again arbitrary.)

Denote

(4) $$\mathcal{P}_m^K := \{P \in \mathcal{P}_m : P \cap K = \phi\} \quad , \quad K \in \mathcal{K} ,$$

(5) $$\mathcal{P}_{mG} := \{P \in \mathcal{P}_m : P \cap G \neq \phi\} \quad , \quad G \in \mathcal{G} .$$

Clearly $\mathcal{P}_m^{\phi} = \mathcal{P}_m$ and $\mathcal{P}_{m\phi} = \phi$.

Take the family

(6) $$\mathcal{S} := \{\mathcal{P}_m^K : K \in \mathcal{K}\} \cup \{\mathcal{P}_{mG} : G \in \mathcal{G}\}$$

and consider it as a subbase of a topology in $\mathcal{P}_m$. (For some basic facts on topology we refer to Kelley [11].)

Denote

(7) $$\mathcal{M} := \{ \bigcap_{i=1}^{r} S_i : S_i \in \mathcal{S} , \; i = 1, 2, \ldots, r \; , \; r < +\infty \}$$

(this is the basis) and

(8) $$\mathcal{F} := \{ \bigcup_{\gamma \in \Gamma} M_\gamma : M_\gamma \in \mathcal{M} \; , \gamma \in \Gamma, \; \Gamma \text{ arbitrary} \} .$$

The elements of $\mathcal{F}$ are called open sets in $\mathcal{P}_m$.
We have a topology in $\mathcal{P}_m$ (This is the topology generated by $\mathcal{S}$, or the "roughest" topology containing $\mathcal{S}$)
Now we have

Theorem 1.1. The function $X : \Omega \to \mathcal{P}_m$ defined by (3) is measurable w.r.t. the Borel σ-algebra defined in $\mathcal{P}_m$ by the topology $\mathcal{F}$. □
The truth of the theorem depends on the following lemma

Lemma 1.2. Let N=(n+1).m and let $\varphi : R^N \to \mathcal{P}_m$ be the following map:

(9) $$\varphi(y) := \{ x \in R^n : \sum_{j=1}^{n} y_{(i-1)(n+1)+j} \, x_j \geq y_{i(n+1)} \, , \, i=1,2,\cdots,m \}, \; y \in R^N.$$

Then the φ is continuous in the topology $\mathcal{F}$. □

Proof. We have to prove that $\varphi^{-1}(\mathcal{P})$ is open in R^N for any $\mathcal{P} \in \mathcal{F}$, where $\varphi^{-1}(\mathcal{P}) \subset R^N$ means the set such that $\varphi(\varphi^{-1}(\mathcal{P})) = \mathcal{P}$, (the "inverse domain" of φ). The definition of $\mathcal{F}$ shows that it is enough to prove that $\varphi^{-1}(\underline{\mathcal{P}}_m^K)$ and $\varphi^{-1}(\underline{\mathcal{P}}_{mG})$ are open for any $K \in \mathcal{K}$ and any $G \in \mathcal{G}$.
Let $y \in \varphi^{-1}(\mathcal{P}_m^K)$ i.e.

(10) $$\varphi(y) \in \mathcal{P}_m \quad \text{and} \quad \overline{\varphi(y)} \cap K = \phi .$$

$\overline{\varphi(y)}$ is a (closed) PS and K is a compact PS. Hence there is a hyperplane $L \subset R^N$ such that φ(y) and K are in two different open halfspaces determined by the L. As K is compact, we can choose L such that it is not parallel to any face of φ(y). Let H be the open halfspace containing $\varphi(y)$. Clearly there is a small neighbourhood $\tau(y)$ of y such that $\varphi(z)\in\mathcal{P}_m$ and $\varphi(z)\in H$ for all $z\in\tau(y)$(here we use the fact that L is not parallel to any face of $\varphi(y)$). But this means that $\varphi(z)\cap K=\phi\ \forall z\in\tau(y)$, that proves the openess of $\varphi^{-1}(\mathcal{P}_m^K)$.

Let $y\in\varphi^{-1}(\mathcal{P}_{mG})$, i.e.

(11) $\qquad \varphi(y)\in\mathcal{P}_m$ and $\varphi(y)\cap G\neq\phi$.

Denote by L_y the affine hull of $\varphi(y)$ (i.e. the translated linear subspace of smallest dimension containing $\varphi(y)$). Denote by $\varphi^o(y)$ the relative interior of $\varphi(y)$ (w.r.t. L_y).

The (11) implies

(12) $\qquad \varphi^o(y)\cap G\neq\phi$,

consequently there is an open ball $B\subset R^n$ such that

(13) $\qquad B\subset G,\quad B\cap L_y\subset\varphi^o(y)$

This shows that there is a neighbourhood $\tau(y)$ of y such that $\varphi(z)\in\mathcal{P}_m$ and

(14) $\qquad \varphi(z)\cap B\neq\phi\ \ \forall z\in\tau(y)$.

This proves the openess of $\varphi^{-1}(\mathcal{P}_{mG})$ and by this the lemma is proved. ∎

<u>Proof of Theorem 1.1:</u>

Denote

(15) $a=(a_{11},a_{12},\ldots,a_{1n},\alpha_1, a_{21},a_{22},\ldots,a_{2n},\alpha_2,\ldots,a_{m1},\ldots a_{mn},\alpha_m)$.

It is easy to see that the map

(16) $a:\Omega\to R^N \quad (N=(n+1)\cdot m)$

is Borel-measurable if all components of a are Borel-measurable.
But

(17) $X(\omega) = \varphi(a(\omega))$

where φ is the map in the Lemma 2.2.
φ is measurable (being continuous) and the composition of two measurable maps is measurable. ■

2. RPS of the second type

The system of inequalities (1) where a_{ij} and α_i are random variables defined on (Ω,Σ,P) is called system of random inequalities (SRI).
Let $A(\omega)$, $\alpha(\omega)$ be as in the previous section and denote

(18) $\Omega(x) := \{\omega\in\Omega : A(\omega)x \geqslant \alpha(\omega)\}$.

We say that x solve SRI with the probability at least p ($0\leq p\leq 1$) if

(19) $P(\Omega(x)) \geqslant p$.

The set

(20) $$V(p) := \{x \in R^n;\ P(\Omega(x)) \geq p\}$$

might be called random polyhedral set of the second type (RPS of the second type).

If all random variables a_{ij}, α_i are discrete, then V(p) may be a PS (i.e. intersection of finite number of closed halfspaces), see, [10] p. 83.

In general we cannot expect that V(p) is a PS. A weaker condition is the convexity of V(p). In fact V(p) is called RPS (of second type) if it is a convex set. Below we shall prove some interesting sufficient conditions for this in the particular case when all a_{ij} are deterministic and only α_i are random variables.

So, let ξ be an m-dimensional random variable defined on the probability space (Ω, Σ, P) and we are asking the convexity of the set

(21) $$V(p) := \{x \in R^n: P(Ax \geq \xi) \geq p\}.$$

Let F_ξ be the distribution function of ξ, hence

(22) $$V(p) = \{x \in R^n: F_\xi(Ax) \geq p\}.$$

By definition a function $\varphi(x)$, $x \in R^k$, is called quasi-concave if all its upper level sets

(23) $$\{x \in R^k: \varphi(x) \geq u\}$$

are convex.

It is clear that if the rang of A is n (let us assume this) then the function $F_\xi(Ax)$ (as a function of $x \in R^n$) is quasi-concave if and only if the function $F_\xi(y)$ is quasi concave on the n-dimensional subspace $L := \{y \in R^m : y = Ax, x \in R^n\}$ of R^m. So we can formulate the following simple

Assertion 2.1. If the distribution function $F_\xi(\underline{y})$, $\underline{y} \in R^m$, of the m-dimensional random variable ξ is quasi-concave on the subspace $L := \{y \in R^m : y = Ax, x \in R^n\}$, i.e. if the sets

(24) $\{ y \in L : F_\xi(y) \geq p \}$

are convex for all $0 \leq p \leq 1$, then the sets V(p) are convex for all $0 \leq p \leq 1$ i.e. the SRI $Ax \geq \xi$ give a RPS of the second type. □

This assertion can be formulated using the probability measure

(25) $\nu_\xi(E) := P(\xi \in E)$, $E \subset R^m$ Borel-measurable, so that

(26) $F_\xi(y) = \nu_\xi(\xi \leq y)$.

In what follows we shall assume that $\nu_\xi(E)$ absolute continuous w.r.t. the Lebesque-measure in R^m, i.e. that $\nu_\xi(E)$ is generated by a density function:

(27) $\nu_\xi(E) = \int_E f_\xi(t)\, dt$, $E \subset R^m$ Lebesque measurable

($\int \cdot\, dt$ means L-integral).

Now

$$(28) \qquad F_\xi(y) = \nu_\xi(\xi \le y) = \int_{t \le y} f_\xi(t)\,dt, \qquad y \in R^m .$$

In the rest of the paper we shall deal with the question: <u>What density functions $f_\xi(t)$ generate quasi-concave distributions $F_\xi(y)$?</u>

Our results are a little more restrictive than needed, because in fact we are looking for $f_\xi(t)$ such that $F_\xi(y)$ is quasi-concave on the subspace L only. The whole method below can be adapted to this case, yielding more general results.

It is convenient to express the quasi-concavity in another way.

One can see easily that the function $\varphi: R^k \to R^1_+$ is quasi-concave if and only if

$$(29) \qquad \varphi(\lambda x + (1-\lambda) y) \geq \min\{\varphi(x), \varphi(y)\} \quad \text{for all } x, y \in R^k \text{ and } 0 \le \lambda \le 1 .$$

For $m=1$, any distribution function $F_\xi(y)$ (define now on R^1) is quasi-concave because it is monotone. For higher dimensions the situation is much more complicated.

A first idea in investigating the quasi concavity of $F_\xi(y)$ (for $m \geq 2$) can be formulated as follows: Is there a quasi-concave density function $f: R^m \to R^1_+$ such that its distribution function F is not quasi-concave? We think the answer is yes, however to find such density functions (even in the case $m=2$) seems to be not an easy task. Hence, the research went into the following direction: Find some well defined subfamilies of quasi-concave density functions f that generate quasi-concave distributions. This direction proved to be fruitful already.

To show the first steps, write analytically the basic question.

Does the condition

(30) $f(\lambda x+(1-\lambda)y) \geq \min\{f(x), f(y)\}$, $x,y \in R^m$, $0 \leq \lambda \leq 1$

implies

(31) $\int\limits_{t \leq \lambda u+(1-\lambda)v} f(t)dt \geq \min\{\int\limits_{x \leq u} f(x)dx, \int\limits_{x \leq v} f(x)dx\}$, $u,v \in R^m$, $0 \leq \lambda \leq 1$?

Let us try to derive (in spite of our doubts) (31) from (30) using a following "trick":

The condition (30) implies

(32) $f(t) \geq \sup\limits_{\lambda x+(1-\lambda)y=t} \min\{f(x), f(y)\}$, $t \in R^m$,

hence

(33) $\int\limits_{t \leq \lambda u+(1-\lambda)v} f(t)dt \geq \int\limits^{*}_{t \leq \lambda u+(1-\lambda)v} \sup\limits_{\lambda x+(1-\lambda)y=t} \min\{f(x), f(y)\}\, dt$,

where $\int^{*}$ is the upper Lebesque-integral, i.e.

$$\int\limits^{*}_{R^m} \psi(t)dt := \inf\{\int\limits_{R^m} \omega(t)dt : \omega(t) \geq \psi(t), t \in R^m, \omega \text{ L-integrable}\}.$$

The upper integral is needed here, because the right hand side of (32) is in general not L-measurable.

If the right hand side of (33) were not less than that of (31) then we were ready.
It is hopeless to prove this (in fact it is in general not true), we try to take on f a more restrictive condition than (30). For this define the following concept: Let a,b≥0, $-\infty<\alpha<+\infty$, $\alpha\neq 0$, $0\le\lambda\le 1$ and denote

(34) $$M_\alpha^{(\lambda)}(a,b)=\begin{cases}0 & \text{if } a\cdot b=0\\ (\lambda a^\alpha+(1-\lambda)b^\alpha)^{1/\alpha} & \text{if } a\cdot b>0\end{cases}$$

("the extended α-means").
For a,b fixed, $M_\alpha^{(\lambda)}(a,b)$ is a non-decreasing function of α (see, e.g. Hardy-Littlewood-Pólya, [12]). For $\alpha=-\infty$, $\alpha=+\infty$ or $\alpha=0$ we define the means by taking limits to get

(35) $$\begin{aligned} M_{-\infty}^{(\lambda)}(a,b) &= \min\{a,b\}\\ M_0^{(\lambda)}(a,b) &= a^\lambda b^{(1-\lambda)}\\ M_{+\infty}^{(\lambda)}(a,b) &= \begin{cases}0 & \text{if } a\cdot b=0\\ \max\{a,b\} & \text{if } a\cdot b>0.\end{cases}\end{aligned}$$

We call the function $\varphi: R^k\to R_+^1$ α-concave, $-\infty\le\alpha\le+\infty$, if

(36) $$\varphi(\lambda x+(1-\lambda)y)\ge M_\alpha^{(\lambda)}(\varphi(x),\varphi(y)) \quad \forall\, x,y\in R^k,\ 0\le\lambda\le 1.$$

The $-\infty$-concave functions are quasi-concave, so the monotony of $M_\alpha^{(\lambda)}$ shows that any α-concave function

is quasi-concave. (The 0-concave function are sometimes called logaithmically concave, log-concave).
Now we have

Theorem 2.1. Let $\alpha \geq -1/m$.
The distribution function of an α-concave density function $f: R^m \to R^1_+$ is $\alpha \cdot (1+m\alpha)^{-1}$-concave, consequently quasi-concave. □

This theorem is a simple consequence of the following integral inequality

Theorem 2.2. Let $f, g : R^m \to R^1_+$ be Lebesque-measurable functions and denote

(37) $$h_\alpha^\lambda(t) := \operatorname*{ess\,sup}_{x \in R^m} M_\alpha^\lambda (f(x/\lambda), g((t-x)/(1-\lambda))), \quad t \in R^m.$$

If $\alpha \geq -1/m$ then

(38) $$\int_{R^m} h_\alpha^{(\lambda)}(t)\,dt \geq M_\beta^{(\lambda)} \left(\int_{R^m} f(x)\,dx, \int_{R^m} g(x)\,dx \right),$$

where $\beta = \frac{\alpha}{1+m\alpha}$. □

Proof of Theorem 2.1. Denote

(40) $$\varphi(x) = \chi_A(x) \cdot f(x), \quad \psi(x) = \chi_B(x) \cdot f(x),$$

where $A = \{t \in R^m : t \leq u\}$, $B = \{t \in R^m : t \leq v\}$.
and χ_A, χ_B are the characteristic functions.
We can write

$$(41)\qquad \int_{t\le \lambda u+(1-\lambda)v} f(t)\,dt = \int_{\lambda A+(1-\lambda)B} f(t)\,dt$$

(where $\lambda A + (1-\lambda)B$ is the algebraic sum of the sets).
The α-concavity of f implies

$$(42)\qquad f(t) \geq \sup_{\lambda x+(1-\lambda)y=t} M_{\alpha}^{(\lambda)}(\varphi(x), \psi(\;)) \qquad t\in R^m.$$

It is clear that the right hand side of (42) is zero if $t\notin \lambda A+(1-\lambda)B$, hence (after writing the right hand side of (42) in the form (37))

$$(43)\qquad \int_{\lambda A+(1-\lambda)B} f(t)\,dt \geq \int_{R^m} \operatorname{ess\,sup}_{x\in R^m} M_{\alpha}^{(\lambda)}(\varphi(x/\lambda), \psi(t-x)/(1-\lambda))\,dt\,.$$

Applying the integral inequality (38) we get the result. ■

For the proof of Theorem 2.2, its sharpenings extensions applications and many other similar results together with a whole history of the integral inequalities of this type, we refer to the papers [15] ÷ [22]
We only note here that for $\alpha>0$ the inequality (38)($m=1$) "almost" coincides with a classical result of Henstock-Macbeath [13] that was extended to higher dimensions by Dinghas [14]. Their paper was almost forgotten and some spacial cases of their inequalities were newly rediscovered nearly 20 years later.

References

1. M.G. Kendall, P.A.P. Moran, "Geometrical Probabilisties", Hafner, New York, 1963.

2. P.A.P. Moran, A note on recent research in geometric probability, J. Appl. Prob., 3(1966), 453-463.

3. P.A.P. Moran, A second note on recent research in geometric probability, Adv. Appl, Prob., 1 (1969), 73-89.

4. G. Matheron, "Random Sets and Integral Geometry", Interscience, New York, 1975.

5. D.G. Kendall, Foundations of a theory of random sets, In: E.F. Harding, D.G. Kendall, Eds., "Stochastic Geometry", Interscience, New York, 1973.

6. L.A. Santaló, "Integral Geometry and Geometric Probability" Addison-Wesley, Reading, Mass., 1976.

7. B.K. Schmidt, T.H. Mattheis, The probability that a random polytope is bounded, Math. of Op. Res., 2 (197), 292-296.

8. A. Rényi, R. Sulanke, Zufällige konvexe Polygone in einem Ringgebiet, Z. Wahrscheinl., 9 (1968), 146-157.

9. W.M. Schmidt, Some results in probabilistic geometry, Z. Wahrscheinl., 9 (1968) 158-162.

10. P. Kall, "Stochastic Linear Programming", Springer, Berlin, New York, 1976.

11. J.L. Kelley, "General Topology", Van Nostrand, New York, 1957.

12. G.H. Hardy, J.E. Littlewood, G. Pólya, "Inequalities" Cambridge Univ. Press, Cambridge, 1951.

13. R. Henstock, A.M. Macbeath, On the measure of sum-sets. (I) The theorems of Brunn, Minkowski and Lusternik, Proc. London Math. Soc., Ser. III. 3 (1953) 182-194.

14. A. Dinghas, Über eine Klasse superadditiver Mengenfunktionale von Brunn-Minkowski-Lusternik-schem Typus Math. Zeitschr. 68 (1957) 111-125.

15. B. Uhrin, "On some inequalities of inverse Hölder type having pplications to tochastic programming problems", Seminar Notes, Mathematics No 2., Hungarian Committee for Applied System Analysis, Budapest, 1975.

16. S. Dancs, B. Uhrin, On a class of integral inequalities and their measure theoretic consequences, J. Math. Anal, Appl., 74 (1980), 388-400.

17. S. Dancs, B. Uhrin, The conditions of equality in some integral inequalities. Publ., Math. Debrecen, 29 (1982), 117-132.

18. B. Uhrin, Convolution of multidimensional unimodal functions, In: Trans. of IX-th Prague Conf. on Int. Th.Decision Functions and Random Processes, Prague, 1982., Vol B., Academia Prague, Reidel, Boston, Mass., 1983, 243-249.

19. B. Uhrin, Extensions and sharpenings of Brunn-Minkowski-Lusternik inequality, Tech. Report No 203, Department of Statistics, Stanford University, Stanford, CA, November 1984.

20. B. Uhrin, Some remarks about the convolution of unimodal functions, Annals of Prob., 12 (1984), 640-645.

21. B. Uhrin, Extensions and sharpenings of Brunn-Minkowski and Bonnesen inequalities, In: K. Böröczky, G. Fejes Tóth, Eds: "Proc. Coll. on Intuitive Geometry", Siófok, Hungary, 1985, Coll. Math. Soc. J. Bolyai, Vol 48, North-Holland, Amsterdam-New York 1986, 929-937.

22. B. Uhrin, A reduction theorem for the measures of sum-sets in R^n, MTA SZTAKI Közlemények, 36/1987, 173-196.

TWO TYPES OF RANDOM POLYHEDRAL SETS

B. Uhrin

Summary

An $m \times n$-matrix $A = \{a_{ij}\}$ and a vector $\alpha = (\alpha_i) \in R^m$ define a polyhedral set $PS := \{x \in R^n : Ax \geq \alpha\}$. One speaks of random polyhedral set (RPS) when a_{ij} and α_i are not real number but random variables. In the literature at least three different types of RPS are defined. The paper presents two of them. The first is when RPS is simply a "PS valued" random variable. It is proved in the paper that a topology in the space of PS-s can be defined so that the "PS-valued" function is continuous, consequently measurable. The second type of RPS discussed in the paper comes when only α is a random (m-dimensional) variable. Here the problem is the convexity of the set $V(p) := \{x \in R^n : P(Ax \geq \alpha) \geq p\}$ for all $0 \leq p \leq 1$, when P is the measure related to the α. It is showed that $V(p)$ is convex for many measures P generated by density functions having some well defined concavity-like properties.

VÉLETLEN POLIHEDRIKUS HALMAZOK KÉT TIPUSÁRÓL

Uhrin Béla

Összefoglaló

Egy A m×n-es mátrix és egy $\alpha \in R^m$ m-dimenziós vektor egy PS := $\{x \in R^n : Ax \geq \alpha\}$ polihedrikus halmazt definiál. Véletlen polihedrikus halmazról /RPS/ akkor beszélünk, amikor az A és α elemei valószinüségi változók. Az irodalomban legalább három különböző típusu RPS van. A cikk ezek közül kettőt tárgyal. Az első típus egyszerüen egy "PS-értékü" valószinüségi változó. A cikkben be van bizonyitva, hogy a PS-ek "terében" vett alkalmas topológiában egy "PS-értékü" leképezés folytonos, tehát mérhető, azaz egy valószinüségi változó. A másik tipus, amalyről a cikkben szó van, akkor fordul elő, amikor csak az α véletlen, de az A nem. Itt a fő probléma a $V(p) := \{x \in R^n : P(Ax \geq \alpha) \geq p\}$ halmaz konvexitása, $0 \leq p \leq 1$-re, ahol P az α-hoz tartozó mérték.
A szerző megmutatja, hogy a V(p) konvex, ha a P-t egy bizonyos konkávitás-szerü tulajdonsággal rendelkező sürüségfüggvény generálja.

\- -

MTA SZTAKI Közlemények 39/1988 p. 151-194.

BRUNN-MINKOWSKI-LUSTERNIK INEQUALITY, ITS SHARPENINGS, EXTENSIONS AND SOME APPLICATIONS

B. UHRIN

Computer and Automation Institute,
Hungarian Academy of Sciences
1502 Budapest, Pf. 63., Hungary

1. Introduction

In what follows denote by $\mathcal{L}_n$ the family of Lebesgue(L)-measurable sets in the n-dimensional Euclidean space R^n, $n\geq 1$, and by μ_n the Lebesgue (L)-measure in R^n. Given two sets $A,B \subset R^n$, A+B means their algebraic (or Minkowski) sum, i.e. the collection of all points of R^n of the form a+b, $a \in A$, $b \in B$.

In general, A+B is not L-measurable for $A,B \in \mathcal{L}_n$ ([1],[2], [3]), hence when speaking of the "measure" of A+B, we think of its inner Lebesgue (L)-measure μ_{n*} defined as

$$\mu_{n*}(C) = \sup \{ \mu_n(A) : A \subseteq C, A \in \mathcal{L}_n \}.$$

Of course, in some particular cases A+B is L-measurable, say, if both sets are compact (consenquently, their sum is compact) or if both sets are convex (their sum is convex). The Brunn-Minkowski-Lusternik (B-M-L)-inequality concerns a lower estimation for $\mu_{n*}(A+B)$ in terms of $\mu_n(A)$ and $\mu_n(B)$. Let us see first the case n=1. If we have two compact sets $A,B \subset R^1$ such that $A \cap B = \{0\}$ (the zero) and A is to the left of 0 and B is to the right of 0, then A+B contains $A \cup B$, hence

(1.1) $$\mu_1(A+B) \geq \mu_1(A) + \mu_1(B).$$

As μ_1 is translation invariant, this inequality holds for any two compact sets. If we have two arbitrary non-empty sets $A,B \in \mathcal{L}_1$ then by the inner regularity of μ_1 we can approximate both A,B from inside by compact sets, $\tilde{A} \subseteq A$, $\tilde{B} \subseteq B$. For the sets $\tilde{A},\tilde{B}$ the inequality (1.1) is valid and this shows that we have

(1.2) $$\mu_{1*}(A+B) \geq \mu_1(A) + \mu_1(B).$$

This inequality is formally true if one of the sets has "infinite measure", implying that their sum has also "infinite measure". To avoid such vague statements, we shall restrict ourselves to sets having finite measures. We see that in the 1-dimensional case the proof of the B-M-L-inequality is quite simple.
In the n-dimensional case one of the proofs goes as follows (see, e.g. Eggleston [4] for details). If A and B are rectangular parallelotopes, i.e.

$$A=\{x \in R^n:\ 0 \leq x_i \leq a_i,\ i=1,2,\dots,n\},\ B=\{x \in R^n:\ 0 \leq x_i \leq b_i,\ i=1,2,\dots,n\},$$

then

(1.3) $$A + B = \{x \in R^n:\ 0 \leq x_i \leq a_i+b_i,\ i=1,2,\dots,n\}$$

and

(1.4) $$\mu_n(A+B) = \prod_{i=1}^{n} (a_i+b_i),\quad \mu_n(A) = \prod_{i=1}^{n} a_i,\quad \mu_n(B) = \prod_{i=1}^{n} b_i.$$

One can see easily that

(1.5) $$\prod_{i=1}^{n}(a_i+b_i) \geq \left(\left(\prod_{i=1}^{n}a_i\right)^{1/n} + \left(\prod_{i=1}^{n}b_i\right)^{1/n}\right)^n$$

and this shows

(1.6) $$\mu_n(A+B)^{1/n} \geq \mu_n(A)^{1/n} + \mu_n(B)^{1/n}.$$

Let

(1.7) $$A = \bigcup_{i=1}^{p} A_i \;, \quad B = \bigcup_{j=1}^{q} B_j \;,$$

where A_i, B_j are parallelotopes such that A_k meets A_ℓ only on boundary (similarly for B_j). For such sets, the inequality (1.6) can be proved by induction on p+q. We cut both A and B by translates of hyperplanes parallel to some R^{n-1}. Let $A^{(1)}$ $A^{(2)}$ be the parts of A being to the "left" and "right" of the haperplane, respectively, and similarly for $B^{(1)}, B^{(2)}$. Let p_1, p_2 (q_1, q_2) be the numbers of parallelotopes unions of which are equal to $A^{(1)}$ and $A^{(2)}$ ($B^{(1)}$ and $B^{(2)}$). Then

(1.8) $$p_1+p_2 \leq p+q-1, \; p_2+q_2 \leq p+q-1$$

and one can choose the hyperplanes such that

(1.9) $$\mu_n(A^{(1)})/\mu_n(B^{(1)}) = \mu_n(A^{(2)})/\mu_n(B^{(2)}) = \omega > 0 .$$

Now, it is clear that

(1.10) $$A + B \supset (A^{(1)} + B^{(1)}) \cup (A^{(2)} + B^{(2)})$$

and the intersection of the two sets on the right hand side of (1.10) has the measure zero.

Applying the induction hypothesis and using (1.9) we get (1.6) for the sets of the form (1.7).

After that standard density considerations shall yield (1.6) for any $A, B \in \mathcal{L}_n$ (of the finite measure): (a) if A,B are compact we approximate them by the sets of the form (1.7); (b) if $A, B \in \mathcal{L}_n$ then using the inner regularity of μ_n we approximate the sets from inside by compact ones (for details see [4]). So we have for any non-empty $A, B \in \mathcal{L}_n$ of finite L-measure

$$(1.11) \qquad \mu_{n*}(A+B)^{1/n} \geq \mu_n(A)^{1/n} + \mu_n(B)^{1/n}$$

This is the B-M-L-inequality.

This inequlaity has been proved first for convex sets by Brunn, [5]. The conditions of equality in (1.11) has been given (for convex sets) first by Minkowski [6], finally Lusternik [7] proved (1.11) (together with the condition of equality) for general measurable sets.

The Lusternik's proofs were not quite correct as it was noticed by Henstock and Macbeath [8]. They gave a first rigorous proof of the conditons of equality in (1.11) (see [8], Theorem 2): If $0 < \mu_n(A), \mu_n(B) < +\infty$ and equality holds in (1.11) then denoting by conv (A) the convex hull of A and $\bar{A} = R^n \setminus A$,

$$(1.12) \qquad \mu_n(\mathrm{conv}(A) \cap \bar{A}) = \mu_n(\mathrm{conv}(B) \cap \bar{B}) = 0,$$

(1.13) conv(A) and conv(B) are homothetic

i.e. there is $\omega > 0$ and $b \in R^n$ such that

(1.13′) $\operatorname{conv}(A) = \omega(\operatorname{conv}(B) + b)$.

Conversely, if (1.12) (1.13) hold then $A+B \in \mathcal{L}_n$ and equality holds in (1.11).

As we have remarked the sum A+B is in general not L-measurable.
Let us write an equivalent definition of A+B

(1.14) $A+B = \{ z \in R^n : A \cap (z-B) \neq \emptyset \}$.

If we write a little more restrictive condition in the brackets of (1.14) we get another set

(1.15) $A \boxplus B := \{ z \in R^n : \mu_n (A \cap (z-B)) > 0 \}$.

Let us call this set "essential sum" of A and B.
One can see easily that the characteristic function, of A+B is equal to

(1.16) $\chi_{A+B}(t) = \sup_{x \in R^n} \chi_A(x) \cdot \chi_B(t-x) \quad , t \in R^n.$

It is also not difficult to see that

(1.17) $\chi_{A \boxplus B}(t) = \operatorname{ess-sup}_{x \in R^n} \chi_A(x) \cdot \chi_B(t-x), \quad t \in R^n,$

where for a non-negative function φ by definition

(1.18) $\operatorname{ess-sup}_{x \in R^n} \varphi(x) := \inf \{ \omega \geq 0 : \varphi(x) \leq \omega$ for a.e. $x \in R^n \}$.

The idea to use "ess-sup" instead of "sup" in the functions of the type (1.17) is due to Dancs (private communication) and has been first studied in Uhrin [9]. Later Brascamp and Lieb [10] also defined the essential sum of two sets. It has been proved, that for $A,B \in \mathcal{L}_n$ the set $A \boxplus B$ is measurable [9], moreover open [10]. (see the Remark 2.2 in the next section). It is clear that

(1.19) $$\mu_{n*}(A+B) \geq \mu_n(A \boxplus B).$$

It turned out that the B-M-L inequality is true also in a following slightly sharper form ([10]).

(1.20) $$\mu_n(A \boxplus B)^{1/n} \geq \mu_n(A)^{1/n} + \mu_n(B)^{1/n}.$$

To keep this paper completely self-contained, let us prove here the inequality (1.20) for n=1.

The proof is based on the already proved ienquality (1.2) and on the concept of density points of a set in R^1. Let us recall (see [8], or Federer [11]) that the point $x \in R^1$ is a density point of $A \in \mathcal{L}_1$ if

(1.21) $$\lim_{\omega \to 0+} \mu_1(A \cap [x-\omega, x+\omega]) / \omega = 2$$

where $[\cdot,\cdot]$ is the closed segment.

Denoting by A^* the set of density points of A it is proved that (see [8],[11])

(1.22) $$\mu_1(A^*) = \mu_1(A).$$

Let $A,B \in \mathcal{L}_1$ be of finite positive measure. Let $a \in A^*$ $b \in B^*$. For sufficiently small $\varepsilon > 0$ we have $\omega > 0$ such that

$$\begin{aligned} &\omega(2-\varepsilon) < \mu_1(A \cap [a-\omega, a+\omega]) \le 2\omega \\ &\omega(2-\varepsilon) < \mu_1(B \cap [b-\omega, b+\omega]) \le 2\omega. \end{aligned} \tag{1.23}$$

Denote $A_1=(A \cap [a-\omega, a))-a$, $A_2=(A \cap (a,a+\omega])-a$
$B_1=b-(B \cap (b,b+\omega])$, $B_2=b-(B \cap [b-\omega, b))$,

where $[\cdot,\cdot)$, $(\cdot,\cdot]$ are half-open intervals.
Clearly $A_1 \cap B_2 = \emptyset$ $A_2 \cap B_1 = \emptyset$.
The conditions (1.23) imply that

$$\mu_1((A_1 \cup A_2) \cap (B_1 \cup B_2)) \geq 2\omega(1-\varepsilon). \tag{1.24}$$

But the left hand side of (1.24) is equal to

$$\mu_1(A_1 \cap B_1) + \mu_1(A_2 \cap B_2), \tag{1.25}$$

consequently one of the terms in (1.25) is positive.
We can see easily that this implies

$$a+b \in A \boxplus B. \tag{1.26}$$

As $a \in A^*$, $b \in B^*$ were arbitrary, we have

$$A \boxplus B \supseteq A^* + B^*. \tag{1.27}$$

Applying now the inequality (1.2) to $A^* + B^*$, taking into account (1.22) we get (1.20) for n=1, i.e.

$$\mu_1(A \boxplus B) \geq \mu_1(A) + \mu_1(B). \tag{1.28}$$

The B-M-L-inequality has been sharpened also in the following way (see Ohmann [12], Bonnesen [13], Bonnesen and Fenchel [14], Hadwiger [15]).

Under the same conditions as in (1.11) we have

$$(1.29) \qquad \mu_{n*}(A+B) \geq \left(m(A)^{1/n-1} + m(B)^{1/n-1}\right)^{n-1} \cdot \left(\frac{\mu_n(A)}{m(A)} + \frac{\mu_n(B)}{m(B)}\right),$$

where m(A) and m(B) are the maximal measures of sections of A and B by translates of R^{n-1} along the n-th coordinate.

One can see easily, using standard inequalities (see e.g. Hardy-Littlewood-Pólya [16]), that the right hand side of (1.29) is not smaller than that of (1.11).
The inequality (1.29) is a special case of the following inequality, where lower dimensional sections are used instead of (n-1)-dimensional ones. Let $S \subseteq R^n$ be a k-dimensional linear subspace, $0 \leq k \leq n$, and let T be an n-k-dimensional complement subspace of S i.e. $S \oplus T = R^n$ (the direct sum). Define for $A \in \mathcal{L}_n$

$$(1.30) \qquad m_k(A) := \operatorname*{ess-sup}_{u \in T} \mu_k(A \cap (S+u)) \quad \text{if } k > 0$$

$$(1.31) \qquad m_0(A) := 1 \quad \text{if } S = \{\Theta\}.$$

Now, for any $0 \leq k \leq n-1$, and $A,B \in \mathcal{L}_n$ such that $0 < m_k(A)$, $m_k(B) < +\infty$, we have (Uhrin [17]):

$$(1.32) \qquad \mu_{n*}(A+B) \geq \left(m_k(A)^{1/k} + m_k(B)^{1/k}\right)^k \cdot \left(\frac{\mu_n(A)}{m_k(A)} + \frac{\mu_n(B)}{m_k(B)}\right).$$

We have to note that the sets $A \cap (S+u)$ are in general not L-measurable in $(S+u)$ (see, e.g. a remark in [8] p. 182), but the Fubini theorem (see, e.g. Bourbaki [18]) shows that $A \cap (S+u)$ is L-measurable for almost all $u \in T$ and

$$(1.33) \qquad \mu_n(A) = \int_T \mu_k(A \cap (S+u))\,du.$$

In what follows by $\int_H \cdot\, dx$ we mean the Lebesgue (L)-integral in the linear subspace $H \subseteq R^n$.
All above mentioned results are simple consequence of the following reduction theorem of Uhrin [19].

<u>Theorem 1.1</u> (see [19], Theorem 1.1).
For any sets $A, B \in \mathcal{L}_n$ of finite positive measure, any k-dimensional linear subspace $S \subseteq R^n$, $0 \le k \le n$ and $0 \le \lambda \le 1$ we have

$$(1.34) \qquad \mu_n(\lambda A \boxplus (1-\lambda)B) \geq (\lambda\, m_k(A)^{1/k} + (1-\lambda)\, m_k(B)^{1/k})^k \cdot \int_0^1 \mu_{n-k}(\lambda A_k(\xi) \boxplus (1-\lambda) B_k(\xi))\, d\xi,$$

where

$$(1.35) \qquad A_k(\xi) := \{u \in T : \mu_k(A \cap (S+u)) \geq m_k(A)\xi\}, \quad 0 \le \xi \le 1,$$

and T is a linear subspace of dimension (n-k) such that $S \oplus T = R^n$ ($B_k(\xi)$ is defined similarly). □

We note that by definition $\mu_0(\theta) = 1$, $\mu_0(\phi) = 0$, where Θ is the zero element of R^n.

In the rest of the paper we shall be dealing with the "convex combination" $\lambda A + (1-\lambda) B$ or "essential convex combination" $\lambda A \boxplus (1-\lambda) B$ of the sets rather than the sum A+B or essential sum $A \boxplus B$. All results concerning convex combinations can be easily transformed to those using sums and vice-versa. It is clear that the inequality (1.34) is from "both sides" sharper than (1.32) in the sense that

$$(1.36) \qquad \mu_{n*}(\lambda A + (1-\lambda) B) \geq \mu_n(\lambda A \boxplus (1-\lambda) B)$$

and

$$(1.37) \qquad \int_0^1 \mu_{n-k}(\lambda A_k(\xi) \boxplus (1-\lambda) B_k(\xi))\, d\xi \geq \lambda^{n-k} \frac{\mu_n(A)}{m_k(A)} + (1-\lambda)^{n-k} \frac{\mu_n(B)}{m_k(B)}$$

(see [19] for details).

The inequality (1.34) can be considered as a tool for getting many new inequalities which are sharper than the known ones (say, we take a sequence of nested subspaces $S_1 \subset S_2 \subset \cdots$ and successively apply the inequality (1.34), see a detailed discussion in [19].) Both the inequality (1.34) and the concepts occuring in it have interesting links with some classical notions of integral geometry (see [19]).

In the rest of this introductory section let us mention a "technical curiosity" proved in [19]. Let $x, y \in R^n_+$

(the orthant of vectors with non-negative components), $0 \leq \lambda \leq 1$, $-\infty \leq \alpha_i \leq +\infty$, $i=1,2,\dots,n$.

Define

$$(1.38) \quad (\lambda x \overset{\alpha}{+} (1-\lambda)y)_i := (\lambda x_i^{\alpha_i} + (1-\lambda) y_i^{\alpha_i})^{1/\alpha_i}, \quad i=1,2,\dots,n,$$

where for $\alpha_i = -\infty, 0, +\infty$, the expressions on the right hand side of (1.38) are defined taking limits to get: $\min\{x_i, y_i\}$, $x_i^{\lambda} \cdot y_i^{(1-\lambda)}$, $\max\{x_i, y_i\}$.

For $A, B \subset R^n_+$ denote

$$(1.39) \quad \lambda A \overset{\alpha}{+} (1-\lambda)B := \{\lambda a \overset{\alpha}{+} (1-\lambda)b : a \in A, b \in B\}$$

If $\alpha_i = 1$ for all i, this sum boils down to usual convex combination of sets and points. Using an analogous proof as for (1.11) we get ([19], Theorem 2.1): If $A, B \subset R^n_+$ are non-empty L-measurable sets of finite measure $0 \leq \alpha_i \leq 1$, $i=1,2,\dots,n$, and $0 < \lambda < 1$, then

$$(1.40) \quad \mu_{n*}(\lambda A \overset{\alpha}{+} (1-\lambda)B)^{\gamma} \geq \lambda \mu_n(A)^{\gamma} + (1-\lambda)\mu_n(B)^{\gamma}$$

where γ is the harmonic mean of α_i-s, $\gamma = (\Sigma \alpha_i^{-1})^{-1}$. For $\alpha_i = 1 \ \forall i$, the inequality (1.40) gives (1.11).

In the following sections we are dealing with integral ineqality extensions of above mentioned inequalities. The main result will be a reduction theorem of which the Theorem 1.1. is a simple particular case.

2. Integral inequalities of Henstock-Macbeath-Dinghas-type

As we have mentioned in the previous section, exact conditions of equality in the B-M-L-inequality has been first formulated and proved by Henstock and Macbeath [8]. Their paper is important also from another point of view. Their proof depended on a one-dimensional integral inequality which later turned to be an important result in itself. Their (or even special cases of their result) has been "rediscovered" in early seventies and used for the proof of some important theorems in stochastic programming, mathematical statistics, probability theory and partial differential equations (for a detailed survey see the last section). The crucial role their paper plays in the field was not quite recognized until the appearence of the paper by Dancs and Uhrin [20], where not only an integral inequality extension of the B-M-L-inequality but also an extension of the sharper ienquality (1.29) have been proved.

Let $\mathfrak{F}_n$ denote the family of non-negative Lebesgue-measurable functions defined on R^n. The Lebesgue(L)-integral of $f \in \mathfrak{F}_n$ is simply written as

$$\int_{R^n} f(x)\,dx. \tag{2.1}$$

We shall use frequently the following well known identity for the integral (2.1)

$$\int_{R^n} f(x)\,dx = \int_0^\infty \mu_n(H(f,\xi))\,d\xi, \tag{2.2}$$

where

(2.3) $H(f,\xi) := \{x \in R^n : f(x) \geq \xi\},\ \xi \geq 0.$

Good references concerning the theory of real functions, measures and integrals are e.g. Dunford-Schwartz [21], Sz-Nagy [22], Bourbaki [18], Dieudonné [23] or Federer [11].

Define for $a,b \geq 0$, $0 \leq \lambda \leq 1$ and $-\infty < \alpha < +\infty, \alpha \neq 0$, the "extended" means as follows

(2.4) $$M_\alpha^{(\lambda)}(a,b) := \begin{cases} 0 & \text{if } a.b=0 \\ (\lambda a^\alpha + (1-\lambda) b^\alpha)^{1/\alpha} & \text{if } a.b>0, \end{cases}$$

and for $\alpha = 0, -\infty, +\infty$,

(2.5) $$M_0^{(\lambda)}(a,b) := \lim_{\alpha \to 0} M_\alpha^{(\lambda)}(a,b) = a^\lambda \cdot b^{(1-\lambda)}$$

(2.6) $$m(a,b) := M_{-\infty}^{(\lambda)}(a,b) := \lim_{\alpha \to -\infty} M_\alpha^{(\lambda)}(a,b) = \min\{a,b\}$$

(2.7) $$M(a,b) := M_{+\infty}^{(\lambda)}(a,b) := \lim_{\alpha \to +\infty} M_\alpha^{(\lambda)}(a,b) = \begin{cases} 0 & \text{if } a.b=0 \\ \max\{a,b\} & \text{if } a.b>0 \end{cases}$$

For basic properites of means we refer to [16] or a recent book by Bullen, Mitrinovic and Vasic [24]. The function $M_\alpha^{(\lambda)}(a,b)$ is for λ, α fixed a positive homogeneous function of (a,b), is increasing in α when other variables are fixed and non-decreasing in (a,b) separately.

The following two inequalities concerning products of means can be checked very easily.

For a,b,c,d ≥ 0, $-\infty \leq \alpha \leq +\infty$, we have

(2.8) $\quad M(a,b)\cdot M_\alpha^{(\lambda)}(c,d) \geq M_\alpha^{(\lambda)}(ac,bd)$,

(2.9) $\quad M_\alpha^{(\lambda)}(a,b)\cdot M_{-\alpha}^{(\lambda)}(c,d) \geq m(ac,bd)$

(These are the simplest cases of a general inequality for product of extended means proved in [17]).
For $f,g\in\mathcal{F}_n$, $-\infty\leq\alpha\leq+\infty$, $0<\lambda<1$ denote

(2.10) $\quad h_\alpha^{(\lambda)}(t) := \operatorname{ess-sup}_{x\in R^n} M_\alpha^{(\lambda)}(f(x/\lambda), g((t-x)/(1-\lambda)))$, $t\in R^n$.

Assertion 2.1. The functions $h_\alpha^{(\lambda)}$ are L-measurable, i.e.

(2.11) $\quad h_\alpha^{(\lambda)} \in \mathcal{F}_n$. $\square$

Proof. We adapt the method from [9], where $h_0^{(\lambda)}\in\mathcal{F}_n$ has been proved. First, let $-\infty<\alpha<+\infty$.

Denote

(2.12) $$f_k(x) = \begin{cases} f(x) & \text{if } f(x)\geq 1/k \\ 1/k & \text{if } f(x) < 1/k \end{cases}$$

(2.13) $$g_k(x) = \begin{cases} g(x) & \text{if } g(x)\geq 1/k \\ 1/k & \text{if } g(x) < 1/k \end{cases} .$$

Then

(2.14) $$M_\alpha^{(\lambda)}(f_k(x/\lambda), g_k((t-x)/(1-\lambda))) =$$
$$= (\lambda f_k(x/\lambda)^\alpha + (1-\lambda)\, g_k((t-x)/(1-\lambda))^\alpha)^{1/\alpha}.$$

The function $(\lambda a^\alpha + (1-\lambda) b^\alpha)^{1/\alpha}$ is a continuous function of $(a,b) \in R^2_+$, the functions $f_k(x/\lambda)$ and $g_k((t-x)/(1-\lambda))$ are L-measurable in $(t,x) \in R^{2n}$ hence the function (2.14) is L-measurable in R^{2n} as a function of (t,x).

The function (2.14) tends monotone decreasingly as $k \to +\infty$ to the function

$$(2.15) \qquad r(t,x) := M_\alpha^{(\lambda)} \left(f(x/\lambda),\ g((t-x)/(1-\lambda))\right)$$

at each point $(t,x) \in R^{2n}$, hence the function $r(t,x)$ is L-measurable in R^{2n}.

The space $L^\infty(R^n)$ (i.e. the space of essentially bounded L-measurable functions) is the dual space to the separable space $L^1(R^n)$ (the space of Lebesgue-intagrable functions). This implies that for $\psi \in L^\infty(R^n)$ we can write

$$(2.16) \qquad \operatorname*{ess-sup}_{x \in R^n} |\psi(x)| = \sup_{\varphi \in D} \int_{R^n} \psi(x) \cdot \varphi(x)\, dx,$$

where D is a countable subset on the unit sphere of the space $L^1(R^n)$ (this is a consequence of the Hahn-Banach theorem). So we can write for any fixed $t \in R^n$

$$(2.17) \qquad h_\alpha^{(\lambda)}(t) = \operatorname*{ess-sup}_{x \in R^n} r(t,x) = \sup_{\varphi \in D} \int_{R^n} r(t,x) \cdot \varphi(x)\, dx .$$

The function $s(t,x) := r(t,x) \cdot \varphi(x)$ belongs to $\mathcal{F}_{2n}$ (because $r(t,x) \in \mathcal{F}_{2n}$), hence by the theorem of Fubini-Tonelli, the function $s(t,x)$ is an L-measurable function of x for almost all t and the function

(2.18) $$s(t) := \int_{R^n} s(t,x)\,dx$$

belongs to $\mathcal{F}_n$. This implies that $h_\alpha^{(\lambda)} \in \mathcal{F}_n$ because it is a supremum of countable many functions from $\mathcal{F}_n$. The functions m(a,b) and M(a,b) can be written as

(2.19) $$m(a,b) = \lim_{k\to\infty} M_{-k}^{(\lambda)}(a,b), \quad M(a,b) = \lim_{k\to+\infty} M_k^{(\lambda)}(a,b),$$

and this shows that the functions $m(f(x/\lambda), g((t-x)/(1-\lambda)))$ and $M(f(x/\lambda), g((t-x)/(1-\lambda)))$ are L-measurable functions of (t,x) in R^{2n}. This implies using above arguments that $h_{-\infty}^{(\lambda)}, h_{+\infty}^{(\lambda)} \in \mathcal{F}_n$. ■

Remark 2.2. The above proof depended on the definition of the ess-sup norm of a function and the duality between L^∞ and L^1, i.e. it uses standard tools of functional analysis. Using deeper results from geometric measure theory, especially the fact $\mu_n(A^* \Delta A) = 0$, wehre A^* is the set of density points of $A \subset R^n$ and Δ means symmetric difference of the sets (see [11], Theorem 2.9.11), Brascamp and Lieb [10] proved a more exact result concerning the measurability of $h_\alpha^{(\lambda)}$. They proved that A⊞B is open for $A,B \in \mathcal{L}_n$ and $h_\alpha^{(\lambda)}$ is lower semicontinuous for $\bar{f}, g \in \mathcal{F}_n$. □

We can extend (2.10) also for $\lambda = 0$ and $\lambda = 1$ taking by definition

(2.20) $$h_\alpha^{(0)}(t) = g(t) \qquad h_\alpha^{(1)}(t) = f(t).$$

Now the proof of the following theorem is very easy.

Theorem 2.3. ([10], Theorem 3.1.) Let $f,g \in \mathcal{F}_1$ be essentially bounded functions and $0 \leq \lambda \leq 1$.

If

$$(2.21) \qquad \underset{x\in R^1}{\text{ess-sup}}\, f(x) = \underset{x\in R^1}{\text{ess-sup}}\, g(x) = \Delta ,$$

then

$$(2.22) \qquad \int_{R^1} h_{-\infty}^{(\lambda)}(t)\, dt \geq \lambda \int_{R^1} f(x)dx + (1-\lambda)\int_{R^1} g(x)dx . \qquad \square$$

<u>Proof</u>. We can assume without loss of generality that $\Delta=1$. Using the notation (2.3) one can see easily that

$$(2.23) \qquad H(h_{-\infty}^{(\lambda)}, \xi) \geq \lambda H(f,\xi) \boxplus (1-\lambda)\, H(g,\xi), \quad 0\leq \xi \leq 1.$$

Taking the measures μ_1 of both sides of (2.23), using the inequality (1. 28), integrating both sides of the resulting inequality over $0 \leq \xi \leq 1$ and taking into account (2.2) we get (2.22). ∎

<u>Remark 2.4.</u> Let us recall here the classical result of Henstock and Macbeath ([8] , Lemma 5, p. 190). They proved for $\alpha > 0$, $0<\lambda<1$, $f,g \in \mathcal{F}_1$,

$$(2.24) \qquad \int_{R^1} \sup_{\lambda x+(1-\lambda)y=t} M_\alpha^{(\lambda)}(f(x),g(y))dt \geq$$

$$\geq \quad M_\alpha^{(\lambda)}(\gamma,\delta)\cdot\left(\lambda \int_{R^1} \frac{f(x)}{\gamma}dx + (1-\lambda)\int_{R^1}\frac{g(x)}{\delta}dx\right)$$

under the assumption that the function in the left hand side integral is L-measurable, where $\gamma = \sup_x f(x)$, $\delta = \sup_x g(x)$. For the proof of (2.24) they used the nice idea

$$(2.23') \qquad H(\tilde{k}_\alpha^{(\lambda)}, \xi) \geq \lambda H(\tilde{f},\xi) + (1-\lambda)\, H(\tilde{g},\xi),\ 0\leq \xi \leq 1,$$

$$\tilde{f} = f/\gamma,\ \tilde{g} = g/\delta\ ,\ \tilde{k}_\alpha^{(\lambda)} = k_\alpha^{(\lambda)} / M_\alpha^{(\lambda)}(\gamma,\delta)\ ,$$

where $k_\alpha^{(\lambda)}$ is the function in the left hand side integral of (2.24), and after that they used the 1-dimensional B-M-L inequality. This idea goes back to Bonnesen's proof of the Brunn-Minkowski inequality (see [13],[14]). It can be find also in [23], p. 238. The same idea is used in [10]. We can state that the inequality (2.22) and the proof of it via (2.23) and (1.28) is principially due to Henstock and Macbeath. □

Using the inequality (2.9) we get from (2.22) immediately

Corollary 2.5. For $f,g \in \mathcal{F}_1$ essentially bounded functions such that

(2.25) $\quad 0<\gamma:=\operatorname{ess-sup}_x f(x), \quad 0<\delta:=\operatorname{ess-sup}_x g(x)$

and for $-\infty \leq \alpha \leq +\infty$, $0\leq\lambda\leq 1$, we have

(2.26) $\quad \int_{R^1} h_\alpha^{(\lambda)}(t)dt \geq M_\alpha^{(\lambda)}(\gamma,\delta)\cdot\left(\lambda\int_{R^1}\frac{f(x)}{\gamma}dx+(1-\lambda)\int_{R^1}\frac{g(x)}{\delta}dx\right).$ □

The inequality (2.26) is trivially true also in the case $\gamma=0$ or $\delta=0$ if we interpret in these cases the right hand side as zero.

Using the Hölder inequality we can easily prove for $\alpha\geq -1$, $0\leq\lambda\leq 1$,

(2.27) $\quad M_\alpha^{(\lambda)}(a,b).M_1^{(\lambda)}(c,d) \geq M_{\alpha/(1+\alpha)}^{(\lambda)}(ac,bd).$

Hence we have a following weakened form of (2.26)

Corollary 2.6. Under the assumptions of the Corollary 2.5, assuming further that $\alpha \geq -1$, we have

$$\int_{R^1} h_\alpha^{(\lambda)}(t)dt \geq M_{\alpha/(1+\alpha)}^{(\lambda)}\left(\int_{R^1} f(x)dx, \int_{R^1} g(x)dx\right). \quad \square \tag{2.28}$$

For $\alpha > 0$ and using "sup" instead of "ess-sup" in the deifnition of $h_\alpha^{(\lambda)}$, this inequality was proved also by Henstock and Macbeath (under the restriction on f and g so that $h_\alpha^{(\lambda)}$ be measurable).

One can see easily that (2.28) holds if we drop the assumption on the boundeness of f and g. Using a simple induction on dimension, the inequality (2.28) can be extended to higher dimensions as well.

Theorem 2.7 ([10],Theorem 3.3). Let $f, g \in \mathcal{F}_n$.
Then for $\alpha \geq -1/n$, $0 \leq \lambda \leq 1$, we have

$$\int_{R^n} h_\alpha^{(\lambda)}(t)dt \geq M_{\alpha/(1+n\alpha)}^{(\lambda)}\left(\int_{R^n} f(x)dx, \int_{R^n} g(x)dx\right). \quad \square \tag{2.29}$$

For $\alpha > 0$, defining $h_\alpha^{(\lambda)}$ with the "sup" and assuming that $h_\alpha^{(\lambda)}$ is L-measurable, the inequality (2.29) is the higher dimensional extension of the result of Henstock and Macbeath due to Dignhas [25] . Letting tend $\alpha \to 0+$ we get from their results an inequality for $\alpha = 0$ (this case has been rediscovered nearly twenty years later, see [26],[27],[28]).

The inequality (2.29) (more precisely its weaker "sup" case) has been proved by some other authors as well, [29],[30].
The first n-dimensional extension of the sharper inequality

(2.26) is due to Dancs and Uhrin [20]. To formulate their result, for $f \in \mathcal{F}_n$ denote

(2.30) $$m_i(f) := \operatorname{ess-sup}_{x_i} \int_{R^{n-1}} f(x)\,dx_1 \ldots dx_{i-1} \cdot dx_{i+1} \cdot \ldots dx_n .$$

Now we have

Theorem 2.8 (see [20], Theorem 3.2 and Remark on p. 398).
Let $f, g \in \mathcal{F}_n$ be such that

(2.31) $$0 < m_i(f) < +\infty , \quad 0 < m_i(g) < +\infty .$$

Let $\alpha \geq -1/(n-1)$, $0 \leq \lambda \leq 1$. Then for $1 \leq i \leq n$ occuring in (2.31) we have

(2.32) $$\int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geq M_\beta^{(\lambda)}(m_i(f), m_i(g)) \cdot$$
$$\cdot \left(\lambda \int_{R^n} \frac{f(x)}{m_i(f)}\,dx + (1-\lambda) \int_{R^n} \frac{g(x)}{m_i(g)}\,dx \right),$$

where $\beta = \alpha / (1 + (n-1)\alpha)$. □

It is clear that $-1/(n-1) < -1/n$ and the right hand side of (2.32) is not less than that of (2.29) i.e. inequality (2.32) sharpens (2.29). For n=1, (2.32) coincides with (2.26).

In [20] a following inequality for $\alpha \leq -1/n$ has been proved.

Theorem 2.9 ([20], Theorem 3.3, Remark p. 398).
Let $f, g \in \mathcal{F}_n$, $-\infty \leq \alpha \leq -1/n$, $0 \leq \lambda \leq 1$.

Then

$$(2.33)\qquad \int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geqslant \min\{\lambda^{n+1/\alpha}\int_{R^n} f(x)\,dx,\ (1-\lambda)^{n+1/\alpha}\int_{R^n} g(x)\,dx\}. \square$$

For $\alpha=-\frac{1}{n}$ the right hand sides of (2.33) and (2.29) coincide. For $-1/(n-1)\leq\alpha<-1/n$ we use the following inequality proved in [17](Lemma 3.1, p.8):

$$(2.34)\qquad M_\alpha^{(\lambda)}(a,b)\cdot M_\beta^{(\lambda)}(c,d) \geqslant \min\{\lambda^\gamma ac,(1-\lambda)^\gamma bd\},$$

where α and β are such that $\alpha+\beta\leq 0$, $\alpha\cdot\beta<0$ and $\gamma:=\frac{\alpha+\beta}{\alpha\beta}$.
The inequality (2.34) shows that for $-1/(n-1)\leq\alpha<-1/n$, the right hand side of (2.32) is not less than that of (2.33) i.e. for such α (2.32) is sharper than (2.33).

We see that inequalities (2.32) and (2.33) improve and extend all inequalities mentioned.

3. A reduction integral inequality

In this section we shall study the following question: What lower estimations can be written for $\int h_\alpha^{(\lambda)}(t)\,dt$ if $\alpha<-1/(n-1)$? We have seen that for such α the inequality (2.33) holds, so the question is whether we can write estimations better than (2.33).

Let $f\in\mathcal{F}_n$ and let $S\subseteq R^n$ be a k-dimensional linear subspace, $0\leq k\leq n$, T its complement subspace, i.e. $S\oplus T = R^n$.

By the theorem of Fubini-Tonelli (see e.g. [21] p.194)

$f(x+u)$ is measurable in $x \in S$ for almost all $u \in T$, the function

$$(3.1) \qquad i(f,u) := \int_S f(x+u)\,dx, \qquad u \in T$$

is measurable in $u \in T$ and

$$(3.2) \qquad \int_{R^n} f(z)\,dz = \int_T i(f,u)\,du.$$

Denote

$$(3.3) \qquad m_k(f) := \operatorname*{ess-sup}_{u \in T} i(f,u), \qquad 0 \leq k \leq n,$$

in particular

$$(3.4) \qquad m_o(f) := \operatorname*{ess-sup}_{x \in R^n} f(x), \quad m_n(f) := \int_{R^n} f(x)\,dx,$$

and

$$(3.5) \qquad \bar{H}(f,\xi) := \{x \in R^n : f(x) \geq m_o(f)\cdot\xi\}, \quad 0 \leq \xi \leq 1.$$

Given two functions $f, g \in \mathcal{F}_n$, denote for $-\infty \leq \alpha \leq +\infty$, $0 < \lambda < 1$,

$$(3.6) \qquad h_\alpha^{(\lambda)}(t) := \operatorname*{ess-sup}_{x \in R^n} M_\alpha^{(\lambda)}(f(x/\lambda), g((t-x)/(1-\lambda))), \quad t \in R^n,$$

and if $0 < m_k(f), m_k(g) < +\infty$ then denote

$$(3.7) \qquad k_\alpha^{(\lambda)}(\tau) := \operatorname*{ess-sup}_{u \in T} M_\alpha^{(\lambda)}\left(\frac{i(f, u/\lambda)}{m_k(f)}, \frac{i(g,(\tau-u)/(1-\lambda))}{m_k(g)}\right), \tau \in T.$$

For $\lambda=0$ or $\lambda=1$ we take by definition

$$(3.8)\qquad h_\alpha^{(0)}(t)=g(t),\ h_\alpha^{(1)}(t)=f(t),\ k_\alpha^{(0)}(\tau)=i(g,\tau)$$

$$k_\alpha^{(1)}(\tau)=i(f,\tau).$$

For $k=n$, i.e. $S=R^n$, $T=\{\theta\}$, we take by definition $k_\alpha^{(\lambda)}(\theta)=1$.

Now, we have

Theorem 3.1. Let $f,g\in\mathcal{F}_n$. Let $S\subseteq R^n$ be a k-dimensional linear subspace $0\leq k\leq n$, and $T\subseteq R^n$ any $(n-k)$-dimensional linear subspace such that $S\oplus T=R^n$. If $0\leq\lambda\leq 1$ and

$$(3.9)\qquad 0<m_0(f),\ m_0(g)<+\infty,$$

then we have

$$(3.10)\qquad \int_{R^n} \operatorname*{ess-sup}_{x\in R^n} \min\left\{\frac{f(x/\lambda)}{m_0(f)},\ \frac{g((t-x)/(1-\lambda))}{m_0(g)}\right\} dt \geq$$

$$\geq \int_0^1 \mu_n(\lambda\bar{H}(f,\xi)\boxplus(1-\lambda)\bar{H}(g,\xi))\,d\xi.$$

If $0<k\leq n$, $\alpha\geq -1/k$, $\beta\geq -\alpha/(1+k\alpha)$, $0\leq\lambda\leq 1$ and

$$(3.11)\qquad 0<m_k(f),\ m_k(g)<+\infty,$$

then

$$(3.12)\qquad \int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geq M_{-\beta}^{(\lambda)}(m_k(f),\ m_k(g))\cdot\int_T k_\omega^{(\lambda)}(\tau)\,d\tau,$$

where $\omega:=(\alpha^{-1}+\beta^{-1}+k)^{-1}$. □

Before the proof let us recall the following inequality concerning products of extended means (see, e.g. [31]). For $a,b,c,d \geq 0$, α, β such that $\alpha+\beta \geq 0$ and $0 \leq \lambda \leq 1$ we have

$$(3.13) \qquad M_{\alpha}^{(\lambda)}(a,b) \cdot M_{\beta}^{(\lambda)}(c,d) \geq M_{\frac{\alpha\beta}{\alpha+\beta}}^{(\lambda)}(ac,bd).$$

Proof of Theorem 3.1: Denote by $h(t)$ the integrand in the left hand side of (3.10). In the previous section we have proved that $h(t)$ is L-measurable (Assertion 2.1). We can see easily that

$$(3.14) \qquad H(h,\xi) \supseteq \lambda \bar{H}(f,\xi) ⊞ (1-\lambda)\bar{H}(g,\xi), \quad 0 \leq \xi \leq 1.$$

Taking the measures of both sides of (3.14) and integrating over $0 \leq \xi \leq 1$ we get via (2.2) the inequality (3.10).

The Assertion 2.1 shows that $h_{\alpha}^{(\lambda)} \in \mathcal{F}_n$ and $k_{\omega}^{(\lambda)} \in \mathcal{F}_{n-k}$. For $k=n=1$ the inequality (3.12) coincides with (2.28). Assume for the moment that (3.12) has been already proved for $k=n-1$, $n > 1$. Then the case $k=n$ can be derived in the following way.

If $k=n-1$, $n > 1$, then T is 1-dimensional and using (3.10), (1.28), (2.2) and (3.2) we have for any ω

$$(3.15) \qquad \int_T k_{\omega}^{(\lambda)}(\tau)\,d\tau \geq \lambda \frac{m_n(f)}{m_{n-1}(f)} + (1-\lambda)\frac{m_n(g)}{m_{n-1}(g)}.$$

Let α be such that

(3.16) $\qquad \alpha \geq -1/n$.

It is clear that for such α, $\alpha \geq -1/(n-1)$, hence we have by (3.12) and (3.15) ((3.12) is true for k=n-1)

(3.17) $$\int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geq M^{(\lambda)}_{\frac{\alpha}{1+(n-1)\alpha}}(m_{n-1}(f), m_{n-1}(g)) \cdot \left(\lambda \frac{m_n(f)}{m_{n-1}(f)} + (1-\lambda)\frac{m_n(g)}{m_{n-1}(g)}\right).$$

The condition $\alpha \geq -1/n \geq -1/(n-1)$ implies that

(3.18) $$\frac{\alpha}{1+(n-1)\alpha} + 1 \geq 0$$

Hence (3.13) applied to the right hand side of (3.17) yield

(3.19) $$\int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geq M^{(\lambda)}_{\frac{\alpha}{1+n\alpha}}(m_n(f), m_n(g)) \geq M^{(\lambda)}_{-\beta}(m_n(f), m_n(g))$$

for any $\beta \geq \frac{-\alpha}{1+n\alpha}$. This proves (3.12) for k=n assuming that it is true for k=n-1.

Now take the pairs (n,k), $1 \leq k \leq n$, into lexicographic order, i.e. $(n_1,k_1) \prec (n_2,k_2)$ if either $n_1 < n_2$ or $\{n_1 = n_2$ and $k_1 < k_2\}$. We get a sequence.

(3.20) $\qquad (1,1) \prec (2,1) \prec (2,2) \prec (3,1) \prec (3,2) \prec (3,3) \prec \ldots$

We proceed with the proof of (3.12) by the induction on this sequence. For n=k=1, (3.12) is true. Assume we have proved (3.12) for all first N-1 members of the sequence (3.20) and let (n,k) be the N-th member of the sequence.

If k=n we are ready by the above reasoning because the case (n, k=n-1) is assumed to be proved by the induction. So let $1 \leq k \leq n-1$, $n > 1$ and

(3.21) $$\alpha \geq -1/k \ , \ \beta \geq \frac{-\alpha}{1+k\alpha} .$$

We can write using (3.13)

(3.22) $$M_\beta^{(\lambda)}(m_k(f)^{-1}, m_k(g)^{-1}) \int_{R^n} h_\alpha^{(\lambda)}(t)\,dt \geq$$

$$\geq \int_T \operatorname{ess-sup}_{u \in T} \left(\int_S \operatorname{ess-sup}_{x \in S} M^{(\lambda)}_{\frac{\alpha\beta}{\alpha+\beta}} \left(\frac{f(\frac{x+u}{\lambda})}{m_k(f)}, \frac{g(\frac{\tau+z-u-x}{1-\lambda})}{m_k(g)} \right) dz \right) d\tau$$

One can easily see that the conditions (3.21) are equivalent to the conditions

(3.23) $$\alpha+\beta \geq 0 \ , \ \frac{\alpha\beta}{\alpha+\beta} \geq -1/k \ ,$$

hence, applying the case (k,k) of (3.12) (which is true by the induction) to the inner integral $\int_S \ldots dz$, we get that the right hand side of (3.22) is not less than

(3.23) $$\int_T \operatorname{ess-sup}_{u \in T} M_\omega^{(\lambda)} \left(\frac{i(f, \frac{u}{\lambda})}{m_k(f)} , \frac{i(g, \frac{\tau-u}{1-\lambda})}{m_k(g)} \right) d\tau. \ \blacksquare$$

The inequalities (3.10) and (3.12) yield automatically results that are sharper than all up to now known inequalities of Henstock-Macbeath-Dinghas-types discussed in the previous section.

Namely, we have

Corollary 3.2 (see [17], Theorem 2.2). For $\alpha \geq -1/k$ denoting $\beta = \frac{\alpha}{1+k\alpha}$, $0 \leq k \leq n$, we have

$$\int_{R^n} h_\alpha^{(\lambda)}(t)dt \geq \tag{3.24}$$

$$\geq \begin{cases} M_\beta^{(\lambda)}(m_k(f), m_k(g))\left(\lambda^{n-k}\frac{m_n(f)}{m_k(f)} + (1-\lambda)^{n-k}\frac{m_n(g)}{m_k(g)}\right) & \text{if } 0 \leq k \leq n-1 \\ M_\beta^{(\lambda)}(m_n(f), m_n(g)) & \text{if } k = n \end{cases} \quad \square$$

Proof. For k=0 apply (2.9) to the left hand side of (3.10), apply (1.37) to the right hand side of (3.10) and integrate over $0 \leq \xi \leq 1$, and after that use (2.2).

For k=n (3.12) coincides with (3.24).

For $0<k\leq n-1$ apply first (3.10) to the integral $\int_T k_\omega^{(\lambda)}(\tau)d\tau$ (the dimension is now n-k), apply again (1.37), integrate over $0 \leq \xi \leq 1$ and use (2.2). ■

The inequality (3.24) is clearly a substantial weakening of (3.10) and (3.12). In spite of this it is sharper than the inequality (2.33). To see this we recall (2.34) i.e. for $\alpha + \beta \leq 0$, $\alpha \cdot \beta < 0$,

$$M_\alpha^{(\lambda)}(a,b) \cdot M_\beta^{(\lambda)}(c,d) \geq \min\{\lambda^{\frac{\alpha+\beta}{\alpha\beta}} ac, (1-\lambda)^{\frac{\alpha+\beta}{\alpha\beta}} bd\}. \tag{3.25}$$

Using these inequalities one can see easily that in each of the domains $-1/k \leq \alpha \leq -1/(k+1)$, k=0,1,2,...,n-2, the inequality (3.24) is sharper than (2.33).
In fact using (3.10) and (3.12) cleverly, we can prove many new and sharp lower bounds for the integral $\int h_\alpha^{(\lambda)}(t)dt$.

4. Measure theoretic consequences of inequalities (3.10) and (3.12)

The reduction Theorem 3.1 yields the Theorem 1.1 at once as a special case. Moreover, it yields easily an extension of Theorem 1.1 to measures other than the Lebesgue measure. Let $f \in \mathcal{F}_n$, $S \subseteq R^n$ a k-dimensional linear subspace, $S \oplus T = R^n$, for $A, B \in \mathcal{L}_n$, χ_A, χ_B mean the characteristic functions of the sets A,B.

Recall the notations (3.1) $\div$ (3.5) and denote

(4.1) $m_k(A) := m_k(\chi_A f)$, $m_k(B) := m_k(\chi_B f)$

(4.2) $A_k(\xi) := \{ u \in T : i(\chi_A \cdot f, u) \geq m_k(A)\xi \}$, $0 \leq \xi \leq 1$,

$B_k(\xi) := \{ u \in T : i(\chi_B \cdot f, u) \geq m_k(B)\xi \}$, $0 \leq \xi \leq 1$.

Denote by ν the measure generated by f, i.e.

(4.3) $\nu(E) := \int_E f(x)\, dx$, $E \in \mathcal{L}_n$.

<u>Theorem 4.1</u> Let $\alpha \geq -1/k$, $0 < \lambda < 1$ and let A,B,f be such that

(4.4) $0 < m_k(A), m_k(B) < +\infty$

and

(4.5) $f(t) \geq \text{ess-sup}_x\, M_\alpha^{(\lambda)}(f(x/\lambda), f(t-x)/(1-\lambda))$

for a.e. $t \in R^n$.

Then denoting $\beta = \frac{\alpha}{1+k\alpha}$ we have

(4.6) $$\nu(\lambda A \boxplus (1-\lambda) B) \geqslant M_{\beta}^{(\lambda)}(m_k(A), m_k(B)).$$

$$\int_0^1 \mu_{n-k}(\lambda A_k(\xi) \boxplus (1-\lambda) B_k(\xi))\,d\xi . \quad \square$$

Proof. After some technical remarks the inequality (4.6) will follow from Theorem 3.1. Denote by $r_{\alpha}^{(\lambda)}(t)$ the right hand side of (4.5) and

(4.7) $$s_{\alpha}^{(\lambda)}(t) := \operatorname*{ess-sup}_{x} M_{\alpha}^{(\lambda)}(\chi_A(x/\lambda)\cdot f(x/\lambda), \chi_B((t-x)/(1-\lambda))\cdot f((t-x)/(1-\lambda))).$$

First we prove that

(4.8) $$\int_{\lambda A \boxplus (1-\lambda) B} r_{\alpha}^{(\lambda)}(t)\,dt = \int_{R^n} s_{\alpha}^{(\lambda)}(t)\,dt.$$

It is clear that

(4.9) $$\chi_{\lambda A \boxplus (1-\lambda) B}(t) = \operatorname*{ess-sup}_{x} M(\chi_A(x/\lambda), \chi_B((t-x)/(1-\lambda))),$$

where M is defined by (2.7).
Using the inequality (2.8) and

(4.10) $$\operatorname*{ess-sup}_{x} \varphi(x)\cdot \operatorname*{ess-sup}_{x} \psi(x) \geqslant \operatorname*{ess-sup}_{x} (\varphi(x)\cdot\psi(x))$$

we see that the left hand side of (4.8) is not less than the right hand side.

On the other hand, if for given t and x

(4.11) $\quad M_{\alpha}^{(\lambda)}(\chi_A(x/\lambda)\cdot f(x/\lambda),\ \chi_B((t-x)/(1-\lambda))\cdot f((t-x)/(1-\lambda)))>0$

then clearly $\chi_A(x/\lambda)=\chi_B((t-x)/(1-\lambda))=1$, i.e.

(4.12) $\quad x\in \lambda A\cap(t-(1-\lambda)B).$

The definition of ess-sup shows that if $s_{\alpha}^{(\lambda)}(t)>0$, then there is a set E such that $\mu_n(E)>0$ and for all $x\in E$ (4.11) holds, i.e. $t\in\lambda A\boxplus(1-\lambda)B$. This shows that the right hand side of (4.8) is not less than the left hand side, consequently (4.8) is true.

Apply now Theorem 3.1 to functions $\chi_A\cdot f$ and $\chi_B\cdot f$. We can easily check that, denoting

(4.13) $\quad C(\xi):=\{\tau\in T:\ k_{\omega}^{(\lambda)}(\tau)\geq\xi\}$

we have

(4.14) $\quad C(\xi)\supseteq\lambda A_k(\xi)\ \boxplus\ (1-\lambda)\ B_k(\xi),\qquad 0\leq\xi\leq 1,$

hence

(4.15) $\quad \int_T k_{\omega}^{(\lambda)}(\tau)\,d\tau\geq\int_0^1\mu_{n-k}\ (\lambda\ A_k(\xi)\boxplus(1-\lambda)\ B_k(\xi))\ d\xi\ .$

By this (4.6) is proved (we apply (3.12) in the sharpest case $\beta=\frac{-\alpha}{1+k\alpha}$). ■

The functions f satisfying (4.5) for some $-\infty\leq\alpha\leq+\infty$ for all $0<\lambda<1$ could be called essentially α-concave. The function $f\equiv 1$ is the "generator" of the Lebesgue measure μ_n, it satisfies (4.5) with $\alpha = +\infty$, hence (4.6) boils down to (1.34). Similarly to Corollary 3.2, a following weaking of (4.6) holds

Corollary 4.2. For $\alpha \geq -1/k$, $0 \leq k \leq n$, denoting $\beta = \frac{\alpha}{1+k\alpha}$ we have

(4.16) $$\nu(\lambda A \boxplus (1-\lambda) B) \geq$$

$$\geq \begin{cases} M_{\beta}^{(\lambda)}(m_k(A), m_k(B)) \cdot (\lambda^{n-k} \frac{\nu(A)}{m_k(A)} + (1-\lambda)^{n-k} \frac{\nu(B)}{m_k(B)}), & \text{if } 0 \leq k \leq n-1 \\ M_{\beta}^{\lambda}(\nu(A), \nu(B)) & \text{if } k=n \end{cases} \quad . \quad \square$$

This inequality sharpen all inequalities of similar type proved in [32], [20], [30], [29].

5. Some applications, remarks

We do not want to give here a survey of applications of B-M-L-inequality. The interested reader can consult e.g. the references [10], [14], [15], [20], [30], [33]. We shall show only three simple examples how the integral inequalities mentioned here can be applied. Chronologically the first is from the field of stochastic programming. Let η be an m-dimensional random variable having the density function $f \in \mathcal{F}_m$ and A be an mxn real-valued matrix, m≥n. A basic problem of so called chance-contrained stochastic programming problem (see e.g. [34]) is:

For what f is the set

(5.1) $\{x \in R^n : \nu(Ax \geqslant \eta) \geqslant p\}$

convex for all $0 \leqslant p \leqslant 1$ where ν is the probability measure generated by f?

This problem was first formulated by Prékopa [35] who later proved that if f is logarithmically concave then (5.1) is convex for all $0 \leqslant p \leqslant 1$, [26]. For the proof he used an integral inequality of Hentock-Macbeath-Dinghas-type: the "sup" case for $\alpha=0$ of the inequality (2.29).

To formulate a more general result, let us call the function $f \in \mathcal{F}_n$ essentially α-concave, $-\infty \leqslant \alpha \leqslant +\infty$, if

(5.2) $$f(t) \geqslant \operatorname*{ess-sup}_{x \in R^n} M_\alpha^{(\lambda)}(f(x/\lambda), f((t-x)/(1-\lambda)))$$

holds for almost all $t \in R^n$ and for all $0 < \lambda < 1$.
Now we have

<u>Theorem 5.1</u> If the density function $f \in \mathcal{F}_m$ of η is essentially α-concave (on R^m) for some $\alpha \geqslant -1/m$, then (5.1) is convex for all $0 \leqslant p \leqslant 1$. □

<u>Proof</u>: Denote by $F(y)$, $y \in R^m$, the distribution function of f, i.e.

(5.3) $$F(y) := \int_{t \leqslant y} f(t)\,dt.$$

For the convexity of (5.1), it is enough to prove that the set

$$(5.4) \qquad \{ y \in R^m : F(y) \geq p \}$$

is convex for all $0 \leq p \leq 1$.

The condition of convexity of (5.4) is clearly equivalent to

$$(5.5) \qquad F(\lambda x + (1-\lambda)y) \geq \min\{F(x), F(y)\}$$

for all $0 \leq \lambda \leq 1$ and $x, y \in R^m$.

But the inequality (5.5) is an easy consequence of the condition (5.2) and the inequality (2.29). ∎

Many densities of mathematical statistics are known to be α-concave for some $\alpha \geq -1/m$ (here " α -concave" means that (5.2) with "sup" instead of "ess sup" holds for all $t \in R^m$), see [32],[33],[34], or a survey in [17]. It is clear that any α -concave function is also essentially α-concave.

Although not explicitly stated in this form, the Theorem 5.1 (or a little weaker version of it with f α- concave) might be a credited to many authors, e.g. [10],[29],[20],[33],[30].

The second example concerns the testing of statistical hypothesis [29].

Let $f \in \mathcal{F}_n$ be a density function, H_o, $A \subset R^n$ be convex sets. Let $bd\,H_o$ means the boundary of H_o. Denote $\Phi(x) = 1 - \chi_A(x)$ where χ_A is the characteristic (indicator) function of A. The question is:
Does the condition

$$(5.6) \qquad \int_{R^n} \Phi(x) \cdot f(x-\eta)\,dx = p \quad \forall\, \eta \in bd\,H_o$$

implies that either (5.6) is true for all $\eta \in H_o$ or

$$(5.7) \quad \begin{cases} \int_{R^n} \Phi(x) \cdot f(x-\eta)\,dx \le p & \forall\, \eta \in H_o \\ \int_{R^n} \Phi(x) f(x-\eta)\,dx \ge p & \forall\, \eta \in R^n \setminus H_o \end{cases} \quad ?$$

In the statistical language: Is it true for f that if Φ is p-similar on $bd\,H_o$ then either Φ is p-similar on H_o or Φ is an unbiased level p test?

Theorem 5.2 If $f \in \mathcal{F}_n$ is essentially α-concave for some $\alpha \ge -1/n$, then for any $0 \le p \le 1$ the answer to the above question is "yes". □

Proof: Using (2.29) and the convexity of A we see that

$$(5.8) \quad \int_{A+\lambda\eta_0+(1-\lambda)\eta_1} f(x)\,dx \ge \min\left\{ \int_{A+\eta_0} f(x)\,dx,\ \int_{A+\eta_1} f(x)\,dx \right\},\ \eta_0, \eta_1 \in H_o,$$

and from this the theorem easily follows. ■

This theorem can be compared with a similar statement in [29], p. 1024.

The last example is of a more analytic nature, although its origin is form the theory of probability (see [31] [36] for details). Let us call the function $f \in \mathcal{F}_n$ essentially α-quasi-concave, $-\infty < \alpha < +\infty$, if

(5.9) $$f(t) \geq \operatorname*{ess-sup}_{x} \min\{\lambda^{\alpha} f(x/\lambda),\ (1-\lambda)^{\alpha} f((t-x)/(1-\lambda))\}$$

hold for a.e. $t \in R^n$ and all $0 < \lambda < 1$.
(Compare with α-quasi-concave functions in [31])
The convolution $f*g$ of two functions $f,g \in \mathcal{F}_n$ is defined as

(5.10) $$f*g(y) := \int_{R^n} f(x)g(y-x)\,dx$$

at each point $y \in R^n$ where the integral exists.

It is known (see, e.g. [18],[21]) that if $\int f(x)\,dx$, $\int g(x)\,dx < +\infty$, then $f*g(y)$ exists for almost all $y \in R^n$ and $\int f*g(y)\,dy < +\infty$
In other words, if $f,g \in L^1(R^n)$ then a "version" of $f*g(y)$ is L-measurable and also belongs to $L^1(R^n)$.
Now, we have

<u>Theroem 5.3</u> (Compare with [31], Theorem 2.1). Let $f,g \in \mathcal{F}_n$ be such that f is essentially α-concave, g essentially β-concave and f,g have finite L-integrals. Then the convolution $f*g$ coincides almost everywhere with a function $\varphi \in \mathcal{F}_n$ having finite integral and such that:

(A) φ is essentially $(\alpha^{-1}+\beta^{-1}+n)^{-1}$ - concave if

$$-1/n \leq \frac{\alpha\beta}{\alpha+\beta} \leq +\infty \ ;$$

(B) φ is essentially $(\alpha^{-1}+\beta^{-1}+n)$-quasi-concave if

$$-\infty \leq \frac{\alpha\beta}{\alpha+\beta} \leq -\frac{1}{n} \qquad \square$$

Proof: The product of essential suprema of two functions is not smaller than the essential supremum of the product of functions. Now, in case (A) we use the inequality (3.13) and after that the inequality (2.29). In case (B) we use first the inequality (2.34) and after that (2.33). ■

Another field, where the integral inequalities of Henstock-Macbeath-Dinghas-type were successfuly used is the theory of diffusion equations [10].

At the end of the paper let s say a few words on the conditions of equalities in integral inequalities mentioned. The proof of conditons (1.12), (1.13) of the equality in (1.11) is quite difficult, hence the proof of such conditions for much more complicated integral ineqaulities is even more difficult.

Up to now two 1-dimensional results have been proved, [37], [38]. To be more specific, let us write the inequality (2.24) and its weakening

(5.11) $$\int_{R^1} k_\alpha^{(\lambda)}(t)dt \geq M_\alpha^{(\lambda)}(\gamma,\delta)\cdot\left(\lambda\int_{R^1}\frac{f(x)}{\gamma}dx + (1-\lambda)\int_{R^1}\frac{g(x)}{\delta}dx\right),$$

(5.12) $$\int_{R^1} k_\alpha^{(\lambda)}(t)dt \geq M_{\frac{1}{1+\alpha}}^{(\lambda)}\left(\int_{R^1} f(x)dx, \int_{R^1} g(x)dx\right),$$

where $f,g \in \mathcal{F}_1$, $0<\gamma=\sup_x f(x)$, $\delta=\sup g(x) <+\infty$, and assuming f and g are such that the function $k_\alpha^{(\lambda)}(t)=\sup_{\lambda x+(1-\lambda)y=t} M_\alpha^{(\lambda)}(f(x),g(y))$ belongs to $\mathcal{F}_1$. Here we assume that $-\infty<\alpha<+\infty$, $0<\lambda<1$ in (5.11) and $-1<\alpha<+\infty$, $0<\lambda<1$ in (5.12).

The conditions of equality in the sharper case (5.11) has been given when f and g are upper semicontinuous [37]. The proof is pretty difficult.

The proof of the conditions for general $f,g\in\mathcal{F}_1$ for the weaker inequality (5.12) is not so difficult already. It depends on a deeper analysis of a proof of (5.12) given in [8] (first proof of Lemma 1, p. 187).

We have performed this analysis in [38], and we have got the following results (see, also [37], p. 131):

Denote $A=\{x\in R^1: f(x)>0\}$, $B=\{x\in R^1: g(x)>0\}$.

Equality is in (5.12) (assuming that $-1<\alpha<+\infty, 0<\lambda<1, 0<\gamma,\delta<+\infty$) if and only if the following conditions are fulfilled:

(5.13) $$A \overset{a.e.}{=} [a_1,a_2] \qquad B \overset{a.e.}{=} [b_1,b_2] \ ;$$

(5.14) $$\left(\frac{\gamma}{\delta}\right)^\alpha = \frac{a_2-a_1}{b_2-b_1} \ ;$$

(5.15) $f(x) = \frac{\gamma}{\delta} g(b_1 + (\frac{\delta}{\gamma})^{\alpha}(x-a_1))$ for a.e. $x \in [a_1, a_2]$,

(5.16) $f(\omega x' + (1-\omega)x'') \geq M_{\alpha}^{(\omega)}(f(x'), f(x''))$ a.e. $x', x'' \in [a_1, a_2]$

where

(5.17) $$\omega = \frac{\lambda \gamma^{\alpha}}{\lambda \gamma^{\alpha} + (1-\lambda)\delta^{\alpha}}$$

Here $[\cdot,\cdot]$ means closed intervals (segments).

References

1. W. Sierpinski, Sur la question de la Mesurabilité de la base de M. Hamel, Fundamenta Math., 1 (1920), 105-111.

2. P. Erdős, A.H. Stone, On the sum of two Borel sets, Proc. Am. Math. Soc., 25 (1970), 304-306.

3. M. Kneser, Summengen in lokalkompakten abelschen Gruppen, Math. Zeitschr., 66 (1956), 88-110.

4. H.G. Eggleston, "Convexity", Cambridge Univ. Press, Cambridge, 1958.

5. H. Brunn, "Über Ovale und Eiflachen", Inaugural Dissertation, München, 1887.

6. H. Minkowski, "Geometrie der Zahlen", Leipzig und Berlin, 1910.

7. L. Lusternik, Die Brunn-Minkowskische Ungleichung für beliebige messbare Mengen, Dokladi Acad. Sci. USSR, 3 (1935), 55-58.

8. R. Henstock, A.M. Macbeath, On the measure of sum-sets. (I) The theorems of Brunn, Minkowski and Lusternik, Proc. London Math. Soc., 3 (1953), 182-194.

9. B. Uhrin, "On some inequalities of inverse Hölder type having applications to stochastic programming problems", Seminar Notes, Mathematics, No. 2, Hungarian Committee for Applied Systems Analysis, Budapest, 1975.

10. H.J. Brascamp, E.H. Lieb, On extensions of Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for logconcave functions and with an application to the diffusion equation, J. of Functional Anal., 22 (1976), 366-389.

11. M. Federer, "Geometric Measure Theory", Springer, Berlin-New York, 1969.

12. D. Ohmann, Eine lineare Verscharfung des Brunn-Minkowskischen Satzes für abgeschlossene Mengen, Arch. der Math., 6 (1955), 33-35.

13. T. Bonnesen, "Les problémes des isopérimetres at des isepiphanes", Paris, 1929.

14. T. Bonnesen, W. Fenchel, "Theorie der konvexe Körper", Springer, Berlin, 1934.

15. H. Hadwinger, "Vorlesungen über Inhalt, Oberfläche und Isoperimetrie", Springer, Berlin, 1957

16. G.H. Hardy, J.E. Littlewood, G. Pólya, "Inequalities", Cambridge Univ. Press, Cambridge, 1951.

17. B. Uhrin, "Sharpenings and extensions of Brunn-Minkowski-Lusternik inequality", Tech. Report No. 203, Dept. of Statistics, Stanford Univ., Stanford, CA, November 1984.

18. N. Bourbaki, "Elements de mathematique: Livre VI. Integration", Hermann, Paris, 1965.

19. B. Uhrin, Extensions and sharpenings of Brunn-Minkowski and Bonnesen inequalities, In: K. Böröczky, G. Fejes Tóth. Eds: "Proc. Coll. on Intuitive Geometry", Siófok, Hungary, 1985, Coll. Math. Soc. J. Bolyai, Vol 48, North-Holland, Amsterdam-New York, 1987, 551-571.

20. S. Dancs, B. Uhrin, On a class of integral inequalities and their measure-theoretic consenquences, J. Math. Anal. Appl., 74 (1980), 388-400.

21. N. Dunford, J.T. Schwartz, "Linear Operators. Part I", Interscience, New York, 1958.

22. B. Sz.-Nagy, "Introduction to real functions and orthogonal expansions", Akadémiai Kiadó, Budapest, 1964.

23. J. Dieudonné, "Foundations of modern analysis. Vol II", Academic Press, New York, 1970.

24. Bullen, D.S. Mitrinovic, Vasic, "Means and their inequalities", Kluwer, Dordrecht, 1988.

25. A. Dinghas, Über eine Klasse superadditiver Mengenfunktionale von Brunn-Minkowski-Lusternik-schem Typus, Math. Zeitschr., 68 (1957), 111-125.

26. A. Prékopa, Logarithmic concave measures with applications to stochastic programming, Acta Sci., Math. (Szeged), 32 (1971), 301-316.

27. L. Leindler, On a certain converse of Hölder's inequality, Acta Sci., Math. (Szeged), 33 (1972), 217-223.

28. A. Prékopa, On logarithmic concave measures and functions, Acta Sci. Math (Szeged), 34 (1973), 335-343.

29. Y. Rinott, On convexity of measures, Ann. of Prob., 4 (1976), 1020-1026.

30. S. Das Gupta, Brunn-Minkowski inequality and its aftermath, J. Multivariate Anal., 10 (1980), 296-318.

31. B. Uhrin, Some remarks about the convolution of unimodal functions, Ann. of Prob., 12 (1984), 640-645.

32. C. Borell, Convex set functions in d-space, Per. Math. Hungar., 6 (1975), 111-136.

33. S. Das Gupta, S-unimodal functions: related inequalities and statistical applications, Sankhya, Series B, 38 (1976) 301-314.

34. P. Kall, "Stochastic linear programming", Springer, Berlin-New York, 1976.

35. A. Prékopa, On probabilistic constrained programming, presentation, Intern. Symp. on Math. Programming, Princeton Univ., Princeton, 1967.

36. B. Uhrin, Convolution of multidimensional unimodal functions, In: Trans. IX-th. Prague Conf. on Inf. Th., Dec. Functions and Random Processes, Prague, 1982, Vol B., Academia, Prague, Reidel, Boston, 1983, 243-249.

37. I. Dancs, B. Uhrin, On the conditions of equality in an integral inequality, Publ. Math. (Debrecen), 29 (1982) 117-132.

38. B. Uhrin, "Systems of linear inequalities, random polyhedral sets and quasi-concave functions" Dissertation, Hungarian Academy of Sciences, Budapest, 1978.

BRUNN-MINKOWSKI-LUSTERNIK INEQUALITY, ITS SHARPENINGS, EXTENSIONS AND SOME APPLICATIONS

B. Uhrin

Summary

Denote by μ_k and μ_{k*} the Lebesgue measure and inner Lebesgue measure, respectively, in R^k, $1 \le k \le n$. For two measurable sets $A, B \subset R^n$, $A+B$ means their algebraic sum. The B-M-L-inequality asserts that $\mu_{n*}(A+B)^{1/n} \ge \mu(A)^{1/n} + \mu(B)^{1/n}$. One of the possible sharpenings of this inequality is:

$$\mu_{n*}(A+B) \ge (m(A)^{1/n-1} + m(B)^{1/n-1})^{n-1} \cdot (\mu_n(A)/m(A)+\mu_n(B)/m(B)),$$

where $m(A)$ and $m(B)$ are the maximal measures of sections of A and B by translates of R^{n-1} along the n-th coordinate (this inequality might be credited to Bonnesen and Ohmann, respectively). Henstock and Macbeath giving a rigorous proof of the conditions of equality in the 1-dimensional B-M-L inequality, formulated on interesting integral inequality extension of the B-M-L inequality, that has been later extended to higher dimensions by Dinghas. The paper deals with developments of the B-M-L inequality along these two directions. The first direction is when the maximal measures of sections of A and B by translates of a k-dimensional linear subspace, $1 \le k \le n-1$, is used (instead of (n-1)-dimensional ones). The second direction is a further development and sharpening of the Henstock - Macbeath-Dinghas (H-M-D) integral inequality, so that to provide integral inequality extensions of measure theoretic inequalities of the first direction. The results of the paper both sharpen and extend all results known up to now. Three interesting applications of the integral inequalities of H-M-D-type are also discussed.

A BRUNN-MINKOWSKI-LUSTERENIK-FÉLE EGYENLŐTLENSÉG ÉLESITÉSEI, KITERJESZTÉSEI, NÉHÁNY ALKALMAZÁSSAL

Uhrin Béla

Összefoglaló

Legyen μ_k ill. μ_{k*} az R^k-ben lévő Lebesgue mérték, ill. belső Lebesgue mérték, $1 \leq k \leq n$. Jelöljük A+B-vel az $A, B \subset R^n$ halmazok algebrai összegét. A $\mu_{n*}(A+B)^{1/n} \geq \mu(A)^{1/n} + \mu(B)^{1/n}$ egyenlőtlenség a B-M-L-egyenlőtlenség. Egy lehetséges élesitése ennek az egyenlőtlenségnek Bonnesen-től ill- Ohmann-tól származik.

Bebizonyitották, hogy $\mu_{n*}(A+B) \geq (m(A)^{1/n-1} + m(B)^{1/n-1})^{n-1}$.

$\left(\frac{\mu_n(A)}{m(A)} + \frac{\mu_n(B)}{m(B)}\right)$, ahol $m(A) := \max\{\mu_{n-1}(A \cap (R^{n-1}+u))\}$ és a maximum az R^{n-1} összes $u \in R^n$ eltoltjára vonatkozik /hasonlóan m(B)/. Henstock ésMacbeath az 1-dmenziós B-M-L-egyenlőtlenség egyenlőségi feltételeinek bizonyitása során megadtak az egyenlőtlenség egy érdekes kiterjesztését /integrál egyenlőtlenséget/, amelyet később Dinghas terjesztett ki magasabb dimenzióra. A cikk a B-M-L egyenlőtlenségnek eme két irányba vett továbbfejlesztésével foglalkozik. Az egyik irány az, amikor a $A \cap (R^{n-1}+u)$ helyett az $A \cap (S+u)$-kat vesszük, ahol S egy k-dimenziós lineáris altér, $1 \leq k \leq n-1$. A másik irány a Henstock-Macbeath-Dinghas (H-M-D)-féle integrál egyenlőtlenség olyan tovább-fejlesztései, illetve élesitései, hogy azok kiadják az első irányban bizonyitott mértékelméleti egyenlőtlenségek "integrál-egyenlőtlenséges" kiterjesztéseit. A cikk eredményei mind kiterjesztik, mind élesitik az eddig ismert eredményeket. A cikk a H-M-D-tipusu egyenlőtlenségeknek három érdekes alkalmazását is tárgyalja.

MTA SZTAKI Közlemények 39/1988 p. 195-228.

DISPERSIONS, PROJECTIONS AND MEASURE OF COVERING IN EUCLIDEAN SPACES

B. UHRIN

Computer and Automation Institute,
Hungarian Academy of Sciences
1502 Budapest, Pf. 63., Hungary

1. Introduction

Let E be a Euclidean space (finite dimensional linear space) and μ be the Lebesgue measure in E. The space E can be considered as a locally compact Abelian group with the vector addition as the group operation and μ as the Haar measure.

For $A,B \subset E$, $A+B$ means the algebraic (Minkowski) addition of the sets, in particular $A+x$ is the translation of A by the x. For any finite set A, $|A|$ is the cardinality of A. For a linear subspace $H \subset E$ let $h_1,\dots,h_k \in H$ be a basis of H, then the set

$$\{x \in H : x = \sum_{i=1}^{k} \lambda_i h_i,\ 0 \le \lambda_i < 1\}$$

is called a basic parallelotope of H. For any set $S \subset E$ by lin(S) we mean the linear hull of S, i.e. the smallest linear subspace containing S. The dimension of any linear subspace $H \subset E$ is denoted by dim (H).

Let $M \subset E$ be any closed (not necessarily discrete) subgroup of E. Given a bounded L-measurable set $A \subset E$, one can ask what portion of E is covered by the family $\{A+m\}_{m\in M}$. In other words, we are interested in evaluation of

$$(1.1) \qquad \mu((A+M) \cap G(r)),$$

where $G(r)$ is the closed Euclidean ball centered at the origin Θ with sufficiently large radius r when compared with the "size" of A.

Let us see two particular cases.

If M is a linear subspace, take a linear subspace T such that $M \oplus T = E$ (the direct sum). Then

$$(1.2) \qquad \mu((A+M) \cap G(r)) \sim \mu(\pi(A) + W) \cdot C(r),$$

where $\pi(A)$ is the projection of the set along M into T, W is a basic parallelotope of M of the unit measure and C(r) is a term depending on r only. One can check easily that the quantity

$$(1.3) \qquad \alpha(M,A) := \mu(\pi(A) + W)$$

depends only on M and A (it does not depend on T and W). We call $\alpha(M,A)$ the measure of covering E by (M-translates of) A.

Let $M \subset E$ be a discrete subgroup and define

$$(1.4) \qquad D(M,A) := \{m\in M:\ m\neq\Theta,\ A \cap (A+m) \neq \emptyset\}.$$

We call this set the dispersion of A (with respect to M).

It is clear that

(1.5) $$D(M,A+x) = D(M,A) \quad \forall\ x\in E.$$

Denote

(1.6) $$M(a) := \{ m\in D(M,A) : a \in A+m \}, \quad a \in A,$$

and

(1.7) $$A(i) := \{ a\in A : |M(a)| = i-1 \},\ i=1,2,\ldots$$

Obviously

(1.8) $$A = \bigcup_{i\geq 1} A(i)$$

and

(1.9) $$\mu(A) = \sum_{i\geq 1} \mu(A(i)).$$

One can easily see that

(1.10) $$\mu((A+M) \cap G(r)) \sim \Big(\sum_{i\geq 1} \frac{\mu(A(i))}{i}\Big) \cdot |M \cap G(r)|$$

We call the quantity

(1.11) $$\alpha(M,A) := \sum_{i\geq 1} \frac{\mu(A(i))}{i}$$

the measure of covering E by (M-translates of) A.

This definition is meaningful also for trivial discrete subgroup $M=\{\Theta\}$, because in this case $D(M,A)=\emptyset$, hence

$A(i)=\emptyset$ for $i>1$, so

(1.11') $\alpha(\{\theta\},A)=\mu(A)$.

If $M\subset E$ is an arbitrary closed subgroup then by a well known theorem (see, e.g. [1]),

(1.12) $M=L\oplus S$,

where L is a discrete subgroup and S a linear subspace. In this case combining (1.3) and (1.11), we shall call the quantity

$$(1.13)\qquad \alpha(M,A):=\sum_{i\geq 1}\frac{\mu((\pi(A))(i)+W)}{i}$$

the measure of covering E by A. Here π is the projection of E along S into $\mathrm{lin}(L)\oplus T$,, where T is a linear subspace such that $\mathrm{lin}(L)\oplus S\oplus T=E$, W is a basic parallelotope of S of unit measure and $(\pi(A))(i)$ is defined as (1.7) with $\mathrm{lin}(L)\oplus T$ instead of E and L instead of M. One can easily see that $\alpha(M,A)$ does not depend on T and W.

The aim of this paper is to prove some interesting properties of $\alpha(M,A)$ especially of the function $\alpha(M,tA)$ for $t\geq 0$. The relation between $\alpha(M_1,A)$ and $\alpha(M_2,A)$, where M_1, M_2 are two subgroups is studied as well.

2. Some general properties of $\alpha(M,A)$

In this section we shall show some general properties of $\alpha(M,A)$ which can be proved quite easily. We shall formulate them in the most abstract and general form, but they can be anderstood better if we identity the quantities and concepts involved using usual "representations" of them.

In what follows in any quotient space, say E/H (where H is a closed subgroup), defined as a collection of different cosets $\{H+x\}, x\in E$, for any element $\omega\in E/H$, (ω) means the subset (H+x) of E. The canonical projection $x \to \{H + x\}$ of E into E/H is denoted by φ if H is a discrete subgroup and by π if H is a linear subspace.

In the case when H is a linear subspace, the E/H can be represented by points of a linear subspace V such that $H \oplus V=E$. If H is discrete subgroup, we take a linear subspace $G \subset E$ such that $\mathrm{lin}(H) \oplus G=E$ and after that take a basic cell Q of H in lin(H). Then E/H can be represented by the points of $Q\oplus G$.

In the case when $M\subset E$ is a closed subgroup, i.e. $M= L\oplus S$, the E/M can be represented by the points of $Q\oplus T$, where T is a linear subspace such that $\mathrm{lin}(L) \oplus S\oplus T =E$ and Q is a basic cell of L in lin(L).

So, let $M\subset E$ be a closed subgroup, $M= L\oplus S$ and $A\subset E$ a bounded L-measurable set.

Assertion 2.1.

(2.1) $\alpha(M,A+x) = \alpha(M,A) \quad \forall\ x\in E,$

(2.1') $\alpha(M,A) \leq \alpha(M,B)$ if $A\subseteq B$. □

Proof: Obvious. ■

Assertion 2.2.

(2.2) $\alpha(M,A) = \tilde{\mu}(\varphi(\pi(A))),$

where, denoting by $\hat{M}\subset E/S$ the discrete subgroup $\{u+S\}_{u\in L}$, φ is the canonical projection of E/S into $(E/S)/\hat{M}$, π is the canonical projection of E into E/S and $\tilde{\mu}$ is the L-measure in E/S. □

Proof: First observe that $(E/S)/\hat{M}$ can be represented by E/M. Identify the spaces involved by their representations: $E/S \sim \text{lin}(L)\oplus T$; $\hat{M}\sim L$; $(E/S)/\hat{M} \sim E/M \sim Q\oplus T$, where Q is a basic cell of L in lin(L); $\varphi: \text{lin}(L)\oplus T \to Q\oplus T$ $\pi: E \to \text{lin}(L)\oplus T$.

Let $\hat{\mu}$ and $\bar{\mu}$ denote the L-measures in $\text{lin}(L)\oplus T$ and S respectively. Then the term $\mu((\pi(A))(i)+W)$ in (1.13) is equal to $\hat{\mu}((\pi(A))(i))$, because $\bar{\mu}(W)=1$ and $\text{lin}(L)\oplus T\oplus S = E$. Hence

(2.3) $$\alpha(M,A) = \sum_{i\geq 1} \frac{\hat{\mu}((\pi(A))(i))}{i} .$$

The right hand side of (2.3) concerns the discrete subgroup L in the space $\mathrm{lin}(L) \oplus T$, so it is nothing else than $\hat{\alpha}(L, \pi(A))$, where $\hat{\alpha}$ is the measure of covering $\mathrm{lin}(L) \oplus T$ by $\pi(A)$. Denote $B = \pi(A)$.

It is clear that

(2.4) $\quad \varphi(\pi(x)) = \varphi(\pi(y))$ if and only if $x-y \in M$

or equivalently

(2.4') $\quad \varphi(\pi(x)) = \varphi(\pi(y))$ if and only if $\pi(x)-\pi(y) \in L$.

On the other hand , one can easily see that $B(i)$ is equal to the union of i-tuples $\{b_1, \ldots, b_i\}$, $b_j \in B(i)$ such that

(2.5) $\quad b_r - b_s \in L$ for all $r, s = 1, 2 \ldots, i$,

and

(2.6) $\quad \{b_1, \ldots, b_i\} \cap \{b'_1, \ldots, b'_i\} = \emptyset$

for any two i-tuples.

Taking the elements of each i-tuple into some prescribed order (say, into lexicographic order), we can collect all first members into one set, all second ones into other set, e.t.c. We get i sets, say $B(i)_j$, $j=1,\ldots,i$, which's union is equal to $B(i)$.
The properties (2.4'), (2.5) and (2.6) show that

(2.7) $\varphi(b_r) = \varphi(b_s)$

for any two members of an i-tuple and

(2.8) $\varphi(b) \neq \varphi(c)$

if and only if b and c belong to different i-tuples.
From these we can easily conclude that

(2.9) $\hat{\alpha}(L,B) = \hat{\mu}(\varphi(B))$

hence (2.2) is true. ■

In fact we have just also proved that

(2.10) $$\alpha(M,A) = \sum_{i\geq 1} \frac{\tilde{\mu}((\tilde{\pi}(A)(i))}{i} = \tilde{\alpha}(\hat{M}, \tilde{\pi}(A))$$

where $(\tilde{\pi}(A))(i)$ concerns the discrete subgroup $\hat{M}$ in E/S and $\tilde{\alpha}$ is the measure of covering in E/S.
If M is discrete then $S=\{\Theta\}$, so $E/S=E$, $\hat{M}=M$, $\tilde{\mu}=\mu$, $\tilde{\pi}(A)=A$.
If M is a linear subspace then $L=\{\Theta\}$, so $\hat{M}=\{\Theta+S\}$ (the trivial "discrete subgroup") hence $D(\hat{M},\tilde{\pi}(A))=\emptyset$, consenquently $\hat{M}(b)=\emptyset$ and $(\tilde{\pi}(A))(i)=\tilde{\pi}(A)$ for i=1 and $(\tilde{\pi}(A))(i)=\emptyset$ for $i\geq 2$, so we have got back to (1.3).
The identity (2.10) shows that

(2.10') $\alpha(M,A) = \alpha(M',A)$

for any two closed subgroups M, M' such that $M=L\oplus S$, $M'=L'\oplus S$ and $\hat{M}=\hat{M}'$.
This fact is clear also from the definition (1.13) of $\alpha(M,A)$.

In the course of the proof we have proved an interesting identity: Let $M \subset E$ be discrete, $A \subset E$ bounded L-measurable and φ the canonical projection $E \to E/M \sim Q \oplus T$.
Then

(2.11) $$\mu(\varphi(A)) = \sum_{i \geq 1} \frac{\mu(A(i))}{i} .$$

We know that

(2.12) $$\mu(A) = \sum_{i \geq 1} \mu(A(i)) .$$

The canonical projection $\varphi(A)$ can be also written as

(2.13) $$\varphi(A) = \{x \in Q \oplus T : A \cap (M+x) \neq \emptyset\}$$

and this implies a following decomposition

(2.14) $$\varphi(A) = \bigcup_{i \geq 1} A_i$$

where

(2.15) $$A_i := \{x \in Q \oplus T : |A \cap (M+x)| = i\}, \; i=0,1,\ldots$$

So we have

(2.16) $$\mu(\varphi(A)) = \sum_{i \geq 1} \mu(A_i) .$$

Using a similar proof as that in the previous assertion, we can prove

(2.17) $$\mu(A) = \sum_{i \geq 1} i \cdot \mu(A_i) .$$

For full dimensional discrete subgroups (lattices) this identity can be found in [2] and played a quite interesting role there. The identity (2.17) is true in more general structures as well, [2].
The identities (2.11) and (2.12) parallels those of (2.16) and (2.17), respectively.

The definition (1.13) of $\alpha(M,A)$ depends only on M and A if E is fixed, but it depends also on E through the measure μ. In many situations $\alpha(M,A)=0$, especially when $\tilde{\pi}(A)$ is of measure zero in $\text{lin}(L)\oplus T$. This is so when, say, A belongs to $E'+x$, where $E'\subset E$ is a linear subspace such that $M\subset E'$. In the latter case, we can measure "the covering of $(E'+x)$ by A" simply taking into (1.13) the L-measure μ' from $(E'+x)$:

$$(2.18) \qquad \alpha'(M,A) = \sum_{i\geq 1} \frac{\mu'((\tilde{\pi}(A))(i) + W)}{i} \quad , \; A\subset E'+x \; .$$

Here $\tilde{\pi}$ and W means the same as in (1.13), i.e. $\tilde{\pi}$ is the projection of E into $\text{lin}(L)\oplus T$ and W is a basic parallelotope of unit volume in S. The $\alpha'(M,A)$ is independent of x in the sense that it is the same for any $(E'+x)$ if the projection of $A\subset(E'+x)$ is the same. More exactly, one can easily see that

$$(2.19) \qquad \mu'((\tilde{\pi}(A))(i) + W) = \mu'((\tilde{\pi}'(A-x)(i) + W)$$

where $\tilde{\pi}'$ is the projection of E' into $\text{lin}(L)\oplus T'$ and T' is such that $\text{lin}(L)\oplus S\oplus T' = E'$.

<u>Assertion 2.3</u> Let $M\subset E'\subset E$, where M is a closed subgroup, E' linear subspace. Let $A\subset E$ be a bounded L-measurable set.

Then

(2.20) $$\alpha(M,A) = \int_{\pi'(A)} \psi(w)\, dw \quad ,$$

where π' is the canonical projection $E \to E/E'$ and

(2.21) $$\psi(w) = \begin{cases} 0 & \text{if } A \cap (w) \text{ is not L-measurable} \\ \alpha'(M, A \cap (w)) & \text{if } A \cap (w) \text{ is L-measurable.} \end{cases}$$

(Recall that (w) means the set $(E' + x)$). □

Roughly speaking (2.20) says that

(2.20') $$\alpha(M,A) = \int_{\pi'(A)} \alpha'(M, A \cap (w))\, dw,$$

if $A \cap (E' + x)$ is measurable for all x (if say, A is convex, then this is certainly true).

<u>Proof:</u> Using (2.10) we transform the whole problem to E/S, E'/S, $\pi(A)$ and $\hat{M}$, so it is enough to prove (2.20) for M discrete.

Let Q be a basic cell of M in lin(M), and T linear subspace such that $\text{lin}(M) \oplus T = E$. Let $T = T_1 \oplus T'$ where $\text{lin}(M) \oplus T_1 = E'$ and $E' \oplus T' = E$. Now we have to prove

(2.22) $$\mu(\varphi(A)) = \int_{\pi'(A)} \psi(y)\, dy$$

where π' is the projection of E into T', φ is the canonical projection of E into $Q \oplus T$ and

(2.23) $$\psi(y) = \begin{cases} 0 & \text{if } A \cap (E' + y) \text{ is not L-measurable} \\ \mu'(\varphi(A \cap (E' + y))) & \text{if } A \cap (E' + y) \text{ is L-measurable.} \end{cases}$$

Writing the identity

(2.24) $$A = \bigcup_{y \in \pi^{-1}(A)} (A \cap (E' + y)$$

we get

(2.25) $$\varphi(A) = \bigcup_{y \in \pi^{-1}(A)} \varphi(A \cap (E' + y)),$$

becauese the sets $A \cap (E' + y)$ are pairwise disjoint.

But

(2.26) $$\varphi(A \cap (E' + y)) \subset E' + y.$$

Indeed, by (2.13) we have

(2.27) $$\varphi(A \cap (E' + y)) = \{x \in Q \oplus T_1 \oplus T' : A \cap (E' + y) \cap (M + x) \neq \emptyset\}$$

and $\mathrm{Lin}(M) \oplus T_1 = E'$, hence for $y \in T'$, $(E' + y) \cap (M + x) = \emptyset$ if $x \notin E' + y$. The inclusion (2.26) implies that

(2.28) $$\varphi(A) \cap (E' + y) = \varphi(A \cap (E' + y))$$

hence by the Fubini theorem (see, e.g. [3]) the function

$\varphi(A) \cap (E'+y)$ is measurable for almost all $y \in \pi'(A)$ and

(2.29) $$\mu(\varphi(A)) = \int_{\pi'(A)} \mu'(\varphi(A) \cap (E'+y))\, dy$$

This gives (2.22) by (2.28). ■

Assertion 2.4. Let $M \subset E$ be a discrete subgroup, $A \subset E$ bounded L-measurable and τ a regular linear transformation of E onto E. Then

(2.30) $$\alpha(M, \tau A) = |\det(\tau)| \cdot \alpha(\tau^{-1} M, A)$$

where τ^{-1} the inverse of τ and $|\det(\tau)|$ is the absolute value of the determinant of τ. □

Proof: The subgroup $M' = \tau^{-1} M$ is also discrete. Using the definitions one can easily see that

(2.31) $$D(M, \tau A) = \tau D(M', A),$$

(2.32) $$|M'(a)| = |M(\tau a)|$$

and

(2.33) $$(\tau A)(i) = \tau((A)'(i)), \quad i=1,2,\dots,$$

where

(2.34) $$(A)'(i) := \{a \in A : |M'(a)| = i-1\}.$$

These give (2.30) at once. ■

<u>Assertion 2.5</u> Let $M = L \oplus S \subset E$ be a closed subgroup, $A \subset E$ bounded L-measurable and τ a linear transformation of E such that

$$(2.35) \qquad \tau S = S \quad \text{and} \quad x \notin S \;\Rightarrow\; \tau x \notin S$$

(i.e. S is the maximal invariant subspace of τ).
Then there is a discrete subgroup $L' \subset E$ such that denoting $M' := L' \oplus S$ we have

$$(2.36) \qquad \tau M' = M$$

$$(2.37) \qquad \alpha(M, \tau A) = |\det(\hat{\tau})| \cdot \alpha(M', A),$$

where $\hat{\tau}: E/S \to E/S$ is the linear regular transformation defined as

$$(2.38) \qquad \hat{\tau}\{x+S\} = \{\tau x + S\}, \quad \{x+S\} \in E/S. \qquad \square$$

<u>Proof:</u> Let $\hat{M}$ and π be as in the Assertion 2.2.
It is clear that $\pi(\tau A) = \hat{\tau}\,\pi(A)$, hence by (2.10) we get

$$(2.39) \qquad \alpha(M, \tau A) = \tilde{\alpha}(\hat{M}, \hat{\tau}\,\pi(A)).$$

The condition (2.35) implies that $\hat{\tau}: E/S \to E/S$ is a linear regular transformation, so applying

(2.30) we get

$$(2.40) \qquad \alpha(M, \tau A) = |\det(\hat{\tau})| \cdot \tilde{\alpha}(\hat{\tau}^{-1}\hat{M}, \pi(A))$$

The set $\hat{\tau}^{-1}\hat{M}$ is a discrete subgroup of E/S, hence it can be written in the form $L' \oplus S$ where L' is a discrete group in E such that $v_1, v_2 \in L'$, $v_1 \neq v_2$ implies $(v_1+S) \cap (v_2 + S) = \emptyset$. Denoting $M'=L' \oplus S$ and using (2.10) we get (2.37). It is clear that for L' (2.36) is true. ■

Remark 2.6. If M is discrete then M' in the previous assertion is unique it equals to $\tau^{-1}M$. If M is a linear subspace, then $L'=\{\theta\}$. For general M, $\hat{\tau}^{-1}\hat{M}$ is unique and there are many discrete subgroups L' "representing" $\hat{\tau}^{-1}\hat{M}$. As to (2.36) it is clear that

$\hat{\tau}(\hat{\tau}^{-1}\hat{M}) = \hat{M}$, hence $\hat{\tau}\{v+S\} = \{\tau v + S\} = \{u+S\}$ for $v \in L'$, $u \in L$. All such L' are good in (2.37) and this implies that $\alpha(M',A)$ does not depend on the particular choice of L'.

(This is clear also from (2.10')). Denote

$$(2.41) \qquad (\tau)^{-1}(x) := \{y \in E : \tau y = x\}$$

(the "inverse domain" of x).

The L' in Assertion 2.5 is any subgroup L' containing exactly one element in each $(\tau)^{-1}(u)$, $u \in L$. □

The last assertion of this section deals with the relation of $\alpha(M_1,A)$ to $\alpha(M_2,A)$, where M_1, M_2 are two closed subgroups of E.

Assertion 2.7. Let $M_1 = L_1 \oplus S_1$ and $M_2 = L_2 \oplus S_2$ be two closed subgroups of E. If

$$(2.42) \qquad S_1 = S_2 \text{ and } \hat{M}_1 \subset \hat{M}_2$$

then

(2.43) $\alpha(M_2,A) \leq \alpha(M_1,A)$.

If (2.42) holds and

(2.44) $D(\widehat{M}_1, \pi_1(A)) = D(\widehat{M}_2, \pi_2(A))$,

then

(2.45) $\alpha(M_2,A) = \alpha(M_1,A)$. □

Proof: The condition $S_1=S_2$ implies that $\pi_1(A)=\pi_2(A)\subset E/S_1$ and $\widehat{M}_1\subset\widehat{M}_2$ are discrete subgroups of E/S_1. Hence by (2.3) it is enough to prove the theorem for $S_1=S_2=\{\Theta\}$, i.e. when $M_1\subset M_2$ are discrete subgroups of E. Denote by $M_1(a)$, $A_1(i)$ and $M_2(a)$, $A_2(i)$ the sets (1.6), (1.7) defined for M_1 and M_2, respectively. The condition $M_1\subset M_2$ implies that $|M_1(a)| \leq |M_2(a)|$ for all $a\in A$, hence

(2.46) $A_1(i) = \bigcup_{j\geq i} (A_1(i) \cap A_2(j))$

and

(2.47) $A_2(i) = \bigcup_{j\leq i} (A_2(i) \cap A_1(j))$.

It is clear that

(2.48) $A = \bigcup_{i\geq 1} A_1(i) = \bigcup_{i\geq 1} A_2(i)$

These three relations imply

(2.49) $\sum_{i\geq 1} \frac{\mu(A_1(i))}{i} \geq \sum_{i\geq 1}\sum_{j\geq i} \frac{\mu(A_1(i)\cap A_2(j))}{j} = \sum_{i\geq 1} \frac{\mu(A_2(i))}{i}$.

This shows (2.43).

If $D(M_1,A)=D(M_2,A)$, then $M_1(a) = M_2(a)$ for all $a \in A$, consenquently $A_1(i) = A_2(i)$ for all i, yielding (2.45). ∎

Remark 2.8. The condition (2.42) is in some sense a necessary condition for the assertion to hold. We shall discuss this question in Section 4. □

3. Discrete subgroups

In this section two main results of the paper will be proved, together with some interesting corollaries.

Theorem 3.1. Let $A \subset E$ be a convex body, and $M \subset E$ a discrete subgroup. Then the function

$$(3.1) \qquad t^{-j}\alpha(M, tA)$$

is non-decreasing in t for $t \geq 0$, where $0 \leq j \leq \dim(E) - \dim(M)$. □

Proof: Because of (2.1) we can assume that $\Theta \in A$. The monotony of (3.1) is clearly equivalent to

$$(3.2) \qquad \alpha(M, \lambda A) \leq \lambda^j \cdot \alpha(M, A) \qquad \forall\ 0 \leq \lambda \leq 1.$$

Hence it is enough to prove (3.2) (or (3.1)) for $j = \dim(E) - \dim(M)$.

Let $E' = V := \mathrm{lin}(M)$ and use the Assertion 2.3. Now $A \cap (V+y)$ is measurable for all $y \in T$, where $V \oplus T \doteq E$, hence

$$(3.3) \qquad \alpha(M, A) = \int_{\pi(A)} \alpha'(M, A \cap (V+y))\, dy \quad ,$$

$$(3.4) \qquad \alpha(M, \lambda A) = \int_{\lambda\pi(A)} \alpha'(M, \lambda A \cap (V+v))\, dv$$

Changing the variable $v = \lambda y$ in the last integral we get

$$(3.5) \qquad \alpha(M, \lambda A) = \lambda^j \int_{\pi(A)} \alpha'(M, \lambda A \cap (V + \lambda y))\, dy \ .$$

Let $y \in \pi(A)$ and $a(y)$ be a point of $A \cap (V+y)$. The convexity of A implies

$$(3.6) \qquad \lambda A \cap (V + \lambda y) - \lambda a(y) \subseteq A \cap (V+y) - a(y) \ .$$

taking into account the definition
(2.18) of $\alpha'(M,A)$ and the remarks after that, the inclusion (3.6) and (2.1') imply

$$\alpha'(M, \lambda A \cap (V+\lambda y)) \leq \alpha'(M, A \cap (V+y)) \tag{3.7}$$

and this proves (3.2). ∎

Let $M_1 \subsetneq M_2 \subsetneq \ldots \subsetneq M_s \subset E$ be discrete subgroups and $A \subset E$ be a convex body. So we have

$$\alpha(M_s, A) \leq \alpha(M_{s-1}, A) \leq \cdots \leq \alpha(M_1, A) \ . \tag{3.8}$$

Define a sequence $\omega_1(A) \leq \omega_2(A) \leq \ \ldots \ \leq \omega_{s-1}(A) \leq \omega_s(A) := +\infty$ as follows

$$\omega_j(A) : \sup\{\omega \geq 0 : \alpha(M_s, tA) = \alpha(M_j, tA) \ \forall\, 0 \leq t \leq \omega\} \ . \tag{3.9}$$

<u>Theorem 3.2.</u> Let $A \subset E$ be a convex body and $M_1 \subsetneq M_2 \ \ldots \subsetneq M_s \subset E$ discrete subgroups. Let $\omega_j \geq 0$, $j=1,2,\ldots,s-1$, be any numbers such that

$$\omega_{j-1} \leq \omega_j \leq \omega_j(A) \ . \tag{3.10}$$

Then

$$\alpha(M_s, tA) \begin{cases} = \alpha(M_1, tA) & \text{if } 0 \leq t \leq \omega_1 \\ \ldots & \\ \geq \left(\frac{t}{\omega_j}\right)^{k_{j+1}} \cdot \left(\frac{\omega_j}{\omega_{j-1}}\right)^{k_j} \cdots \left(\frac{\omega_2}{\omega_1}\right)^{k_2} \cdot \alpha(M_1, \omega_1 A) & \text{if } \omega_j \leq t \leq \omega_{j+1} \\ & j=1,\ldots,s-1,\ \omega_s := +\infty , \end{cases} \tag{3.11}$$

where $k_j = \dim(E) - \dim(M_j)$, $j=1,2,\ldots,s$. □

Proof: The theorem is a simple consequence of Theorem 3.1. For $0 \le t \le \omega_1$, $\alpha(M_s, tA) = \alpha(M_1, tA)$ by the definition of $\omega_1(A)$. Assume we have proved (3.11) for $t = \omega_j$ and let $\omega_j \le t \le \omega_{j+1}$. By the definition of $\omega_{j+1}(A)$ we have

$$t^{-k_{j+1}}\alpha(M_s, tA) = t^{-k_{j+1}}\alpha(M_{j+1}, tA) \tag{3.12}$$

and using the monotony of the right hand side of (3.12) we get

$$t^{-k_{j+1}}\alpha(M_s, tA) \ge \omega_j^{-k_{j+1}}\alpha(M_{j+1}, \omega_j A)\,. \tag{3.13}$$

But

$$\alpha(M_{j+1}, \omega_j A) = \alpha(M_s, \omega_j A) \tag{3.14}$$

and applying the proved inequality for $t = \omega_j$ we get (3.11) ■

Remark 3.3. We have no assumptions on the dimensions of M_j. If, say, $k_j = k_{j+1}$, i.e. $M_j \subset M_{j+1}$ but $\mathrm{lin}(M_j) = \mathrm{lin}(M_{j+1})$, or $M_j = M_{j+1}$, then

$$\alpha(M_s, tA) \ge \left(\frac{t}{\omega_{j-1}}\right)^{k_j}\left(\frac{\omega_{j-1}}{\omega_{j-2}}\right)^{k_{j-1}}\cdots\left(\frac{\omega_2}{\omega_1}\right)^{k_2}\alpha(M_1, \omega_1 A) \quad \text{if } \omega_{j-1} \le t \le \omega_{j+1} \tag{3.15}$$

and in this case ω_j disappears from (3.11), i.e. the result is the same as that without M_j.

This shows that in spite of the fact that there is no bound on the number s of subgroups, the theorem is effective only if all k_j-s are different and in this case $s \le \dim(E)$ (or $s \le \dim(E)+1$ if $\dim(M_1) = 0$). □

Corollary 3.4 Let $A \subset E$ and $M_1 \subseteq \ldots \subseteq M_s \subset E$ be as in Theorem 3.2. Let $0 \le \mu_1 \le \mu_2 \le \cdots \le \mu_{s-1} \le \mu_s := +\infty$ be such that

$$D(M_s, \mu_j A) = D(M_j, \mu_j A) \quad , \quad j=1,2,\ldots,s-1. \tag{3.16}$$

Then

$$\alpha(M_s, tA) \begin{cases} = \alpha(M_1, tA) & \text{if } 0 \le t \le \mu_1 \\ \ge \left(\frac{t}{\mu_j}\right)^{k_{j+1}} \left(\frac{\mu_j}{\mu_{j-1}}\right)^{k_j} \cdots \left(\frac{\mu_2}{\mu_1}\right)^{k_2} \alpha(M_1, \mu_1 A) & \text{if } \mu_j \le t \le \mu_{j+1} \\ & j=1,2,\ldots,s-1 \end{cases} \tag{3.17}$$

where $k_j = \dim(E) - \dim(M_j)$. □

Proof. By the Assertion 2.7, the condition (3.16) implies that

$$\alpha(M_s, \mu_j A) = \alpha(M_j, \mu_j A) \tag{3.18}$$

Hence $\mu_j \le \omega_j(A)$, so the conditions of Theorem 3.2 are satisfied for $\omega_j = \mu_j$. ■

Lemma 3.5. Let $M \subset E$ be a discrete subgroup and $A \subset E$ bounded L-measurable set such that

$$\dim\,(D(M,A) \cup \{\Theta\}) = k \tag{3.19}$$

Then there is a discrete subgroup $L \subseteq M$ such that $\dim\,(L) = k$ and

$$D(L,A) = D(M,A) \ . \quad \square \tag{3.20}$$

Proof. Use the obvious identity

(3.21) $\{\Theta\} \cup D(M,A) = ((A-A) \cap M)$.

We see that $\dim((A-A) \cap M)=k$. Take $S=\mathrm{lin}((A-A) \cap M)$. Then $L:= M \cap S$ is the wanted subgroup, because $(A-A) \cap L=(A-A) \cap M \cap S= = (A-A) \cap M$. ∎

Corollary 3.6. Let $A \subset E$ be a convex body and $M \subset E$ a discrete subgroup. Let $0 \leq \lambda_1 \leq \cdots \leq \lambda_{s-1} \leq \lambda_s := +\infty$ be such numbers that

(3.22) $$\dim(D(M, \lambda_j \cdot A) \cup \{\Theta\}) = \dim(E) - k_j, \ j=1,2,\dots,s-1,$$

where $k_1 \geq k_2 \geq \dots \geq k_{s-1}$, $k_s := \dim(E) - \dim(M)$.

Then

(3.23) $$\alpha(M,tA) \begin{cases} = \alpha(L,tA) & \text{if } 0 \leq t \leq \lambda_1 \\ \geq \left(\frac{t}{\lambda_j}\right)^{k_{j+1}} \left(\frac{\lambda_j}{\lambda_{j-1}}\right)^{k_j} \cdots \left(\frac{\lambda_2}{\lambda_1}\right)^{k_2} \alpha(L, \lambda_1 A) & \text{if } \lambda_j \leq t \leq \lambda_{j+1} \\ & j = 1,2,\dots,s-1, \end{cases}$$

where $L:= M \cap \mathrm{lin}(\lambda_1(A-A) \cap M)$. □

Proof: By the Lemma 3.5 there is a sequence of subgroups

$$M_1 \subseteq M_2 \subseteq \dots \subseteq M_s = M \text{ such that}$$

(3.24) $$\dim(M_j) = \dim(D(M, \lambda_j \cdot A) \cup \{\Theta\}) = \dim(E) - k_j$$

and

(3.25) $$D(M_j, \lambda_j A) = D(M, \lambda_j A)).$$

Indeed, $M_j := M \cap \operatorname{lin}(\lambda_j(A-A)\cap M)$, $j = 1,2,\dots,s-1$, and $\lambda_j(A-A)\cap M \subseteq \lambda_{j+1}(A-A)\cap M$ implies that $M_j \subseteq M_{j+1}$. Let $L := M_1$ and apply the Corollary 3.4. ∎

Let $\Lambda \subset E$ be a discrete subgroup of full dimension, $\dim(\Lambda) = \dim(E)$ (a point lattice) and let $A \subset E$ be a convex body. Denote by ν_j the successive minima of A-A, i.e.

$$(3.26) \qquad \nu_j := \inf\{t \geq 0 : \dim(t(A-A)\cap\Lambda) \geq j\},\ j = 1,2,\dots,n = \dim(E).$$

Then we have

<u>Corollary 3.6.</u>

$$(3.27) \qquad \alpha(\Lambda, tA) \begin{cases} = t^n \mu(A) & \text{if } 0 \leq t \leq \nu_1 \\ \geq t^{n-j}\left(\prod_{i=1}^{j} \nu_i\right)\cdot\mu(A) & \text{if } \nu_j \leq t \leq \nu_{j+1} \end{cases}$$

$$j = 1,2,\dots,n,\ \nu_{n+1} = +\infty$$

<u>Proof:</u> Assume first that

$$(3.28) \qquad 0 < \nu_1 < \nu_2 < \dots < \nu_n .$$

Then for sufficiently small $\varepsilon > 0$ denoting $\bar{\nu}_i = \nu_i - \varepsilon$ we have $\nu_{i-1} < \bar{\nu}_i$, $\bar{\nu}_1 < \bar{\nu}_2 < \dots < \bar{\nu}_n$, consequently by the definition of ν_i

$$(3.29) \qquad \dim(\bar{\nu}_i(A-A)\cap\Lambda) = i-1, \quad i = 1,2,\dots,n.$$

Applying the Corollary 3.6 and taking into account that $\alpha(\{\theta\}, A) = \mu(A)$ we get

$$(3.30)\qquad \alpha(\Lambda, tA)\begin{cases} = t^n\mu(A) & \text{if } 0 \le t \le \bar{\nu}_1 \\ \ge t^{n-j}\left(\prod_{i=1}^{j}\bar{\nu}_i\right)\mu(A) & \text{if } \bar{\nu}_j \le t \le \bar{\nu}_{j+1} \\ & j=1,\dots,n. \end{cases}$$

If, say, $\nu_{j-1} < \nu_j = \nu_{j+1} < \nu_{j+2}$, then denoting again $\bar{\nu}_j = \nu_j - \varepsilon$, $\dim(\bar{\nu}_j(A-A)\cap\Lambda) = \dim(\bar{\nu}_{j+1}(A-A)\cap\Lambda) = j-1$, and, by the Remark 3.3, (3.30) holds again. Letting tend $\varepsilon \to 0$ we get from (3.30) the estimation (3.27). ■

Applying (3.27) for j=n and taking into account (2.2) we get

$$(3.31)\qquad \mu(\varphi(tA)) \ge \left(\prod_{i=1}^{n}\nu_i\right)\mu(A) \qquad \text{if } t \ge \nu_n .$$

This is an inequality proved earlier (see, e.g. [4]), and this gives the successive minima theorem at once ([4]). For symmetric convex bodies (i.e. such that A=-A), the inequality (3.27) (where $\mu(\varphi(tA))$ is instead of $\alpha(\Lambda, tA)$) can be found in [5].

4. Arbitrary closed subgroups

The question is which results of the previous section can be extended to the case when $M = L \oplus S$ is an arbitrary closed subgroup. It turned out that the basic Theorem 3.1 holds true also in this general case.

Theorem 4.1. Let $A \subset E$ be a convex body and $M = L \oplus S \subset E$ a closed subgroup. Then the function

$$(4.1) \qquad t^{-j} \alpha(M, tA)$$

is non-decreasing in t for $t \geq 0$, where $0 \leq j \leq \dim(E) - \dim(M)$. □

Proof: The same as that of Theorem 3.1 . ■

The extension of the remaining parts of the results for discrete subgroups to general closed subgroups seems to be quite a difficult question: it depends on the validity of $\alpha(M_2, A) \leq \alpha(M_1, A)$. The following examples show that this inequality is in general not true if $S_1 \neq S_2$.

Let $e_1, e_2, e_3 \in R^3$ be the usual unit vectors and let $L_1 := \text{int}(e_1, e_2)$ (int (.) means the "integer" hull of the vectors, i.e. the point lattice generated by them), $L_2 := \text{int}(e_1)$ $S_1 = \text{lin}(e_3)$, $S_2 := \text{lin}(e_2, e_3)$. It is clear that $M_1 := L_1 \oplus S_1 \subset M_2 := L_2 \oplus S_2$. Now $E/S_1 \sim \text{lin}(L_1)$, $E/S_2 \sim \text{lin}(L_2)$, $\hat{M}_1 \sim L_1$, $\hat{M}_2 \sim L_2$, hence $\hat{M}_2 \subset \hat{M}_1$, implying $\alpha(M_1, A) \leq \alpha(M_2 A)$ for all $A \subset R^3$. Let S_1 and M_2 be as previously and $L_1' = \text{int}(a)$, where $a = (1,1,0)$. Then $E/S_1 \sim \text{lin}(L_2) \oplus \text{lin}(e_2)$, $L_1' \subset L_2 \oplus \text{lin}(e_2)$ One can easily find sets $A, B \subset R^3$ such that $0 < \alpha(M_1', A) < \alpha(M_2, A)$

and $\alpha(M_1',B) > \alpha(M_2,B) > 0$, where $M_1' = L_1' \oplus S_1$. Let M_2 be again as above and L_1'' be a discrete subgroup of $\mathrm{lin}(e_2)$. Let $M_1'' = L_1'' \oplus \mathrm{lin}(e_2)$. Then there are $A,B \subset R^3$ such that $\alpha(M_1'',A) < \alpha(M_2,A)$ and $\alpha(M_1'',B) > \alpha(M_2,B)$. Finally the same is true if $\tilde{L}_1 \subset L_2$ i.e. denoting $\tilde{M}_1 = \tilde{L}_1 \oplus S_1$, M_2 as above, one can find sets $A,B \subset R^3$ s.t. $\alpha(\tilde{M}_1,A) < \alpha(M_2,A)$ and $\alpha(\tilde{M}_1,B) > \alpha(M_2,B)$. In all four above cases $M_1 \subset M_2 \subset R^3$, but $\alpha(M_2,A) \leq \alpha(M_1,A)$ is not true for all A. The reason for this is that $S_1 \neq S_2$. Using the Assertion 2.7 only a following extension of the Theorem 3.2 can be proved.

<u>Theorem 4.2.</u> Let $A \subset E$ be a convex body and $M_i := L_i \oplus S \subset E$ $i=1,2,\ldots,s$, be closed subgroups such that

(4.2) $$\hat{M}_1 \subset \hat{M}_2 \subset \ldots \subset \hat{M}_s \subset E/S.$$

Let $\omega_j \geq 0$, $j=1,2,\ldots,s-1$, be numbers satisfying

(4.3) $$\omega_{j-1} \leq \omega_j \leq \omega_j(A),$$

where $\omega_j(A)$ are defined by (3.9). Then

(4.4) $$\alpha(M_s,tA) \begin{cases} = \alpha(M_1,tA) & \text{if } 0 \leq t \leq \omega_1 \\ \geq \left(\frac{t}{\omega_j}\right)^{k_{j+1}} \left(\frac{\omega_j}{\omega_{j-1}}\right)^{k_j} \cdots \left(\frac{\omega_2}{\omega_1}\right)^{k_2} \alpha(M_1,\omega_1 A) & \text{if } \omega_j \leq t \leq \omega_{j+1} \end{cases}$$

$$j=1,2,\ldots,s-1,\ \omega_s := +\infty,$$

where $k_j = \dim(E) - \dim(M_j)$ $j=1,2,\ldots,s$. □

It is clear that all results of the previous section can be written down for closed subgroups M_i fulfilling the assumptions of the previous theorem.

5. Concluding remarks

1. Let $\Lambda \subset E$ be a point lattice (discrete subgroup of full dimension). Using the identity (2.2) a result in [2] yields

$$\alpha(\Lambda, A) \geq 2\mu(A)\,(2 + |D(\Lambda, A)|)^{-1}. \tag{5.1}$$

Even some sharper estimations of similar type (using again (2.2)) can be proved (see [2], [6] for details).

In [7] we have proved a converse inequality

$$\alpha(\Lambda, A) \leq (\mu((A-A)\cap\Lambda; A) - \mu(A)) \cdot |D(\Lambda, A)|^{-1}, \tag{5.2}$$

where $\mu(B; A)$, $B \subseteq \Lambda$, is a sort of "measure of A".
The inequality (5.2) can be weakened to yield

$$\alpha(\Lambda, A) \leq \frac{\mu(A-A+A) - \mu(A)}{|D(\Lambda, A)|} \tag{5.3}$$

(an inequality essentially due to H. Hadwiger, [8], via (2.2), see [7])
The inequality (5.1) can be extended to more general structures as well, [2].
The proof of (5.1) depended on the identities (2.16) and (2.17), hence using the same method we get easily that

$$\alpha(M, A) \geq 2\mu(A)\cdot(2 + |D(M, A)|)^{-1} \tag{5.4}$$

where now $M \subset E$ is any discrete subgroup.

Similarly, (5.2) is true for lower dimensional discrete subgroups M as well.

(5.5) $\alpha(M,A) \le (\mu((A-A)\cap M;A) - \mu(A)) \cdot |D(M,A)|^{-1}$

(for the exact meaning of μ(B;A) see [7]).

2. Let $\Lambda \subset E$ be a point lattice again. Using the identity (2.2) the results in [5] and [4] give

(5.6) $\alpha(\Lambda, tK) \begin{cases} = t^n \mu(K) & \text{if } 0 \le t \le \lambda_1/2 \\ \ge t^{n-j} \mu(K) \prod_{i=1}^{j} (\lambda_i/2) & \text{if } \lambda_j/2 \le t \le \lambda_{j+1}/2 \end{cases}$

$j = 1, 2, \ldots, n$,

(5.7) $\alpha(\Lambda, \nu_n C) \ge \mu(C) \prod_{i=1}^{n} \nu_i$

where K is a symmetric convex body (i.e.K=-K), λ_i successive minima of K, C a convex body and ν_i successive minima of C-C.
As we have seen, both result are special cases of our Corollary 3.6.
Both inequalities (5.6) and (5.7) has been used for the proof of successive minima theorem.

A decomposition of A of the type (1.8) and a sum of the type on the right hand side of (1.11) can be found in [9]. They played a role in one of the proofs of successive minima theorem.

We can state that the research reported on in this paper stem from two sources: the sharpening and extension (5.1) of Minkowski convex body theorem, [10], [11], and a detailed study of successive minima theorem .

3. As to extensions of the results to sequences of closed

subgroups $M_i := L_i \oplus S_i$ such that S_i are not equal each to other, we have some preliminary results but they are formally so complicated that it is quite difficult to see what is going on.

The result (5.1) can be extended to structures more general than E: locally compact Abelian groups and their discrete subgroups (see, [2]).

An interesting theme of study would be to extend the results of this paper (at least those in Section 3) to these general structures. This seems to be not hopeless.

References

1. E. Hewitt, K.A. Ross, "Abstract Harmonic Analysis", Vol I., Springer, Berlin-New York, 1963.

2. B. Uhrin, Some remarks on the lattice points in difference sets, In: J. Szabados, K. Tandori, Eds.: "Proc. of the A. Haar Memorial Conference", Budapest, 1985. Coll. Math. Soc. J. Bolyai, Vol. 49, North-Holland, Amsterdam-New York, 1987, 929-937.

3. N. Bourbaki, "Elements de Mathematique. Livre VI. Integration", Hermann, Paris, (Russian translation, Nauka, Moscow, 1967).

4. P.M. Gruber, C.G. Lekkerkerker, "Geometry of Numbers", 2nd Edition, North-Holland, Amsterdam-New York, 1987.

5. J.W.S. Cassels, "An Introduction to Geometry of Numbers" Springer, Berlin-New York, 1959.

6. G. Freiman, A. Heppes, B. Uhrin, A lower estimation for the cardinality of finite difference sets in R^n, In: Coll. Math, Soc. J. Bolyai, North-Holland, Amsterdam-New York, to appear

7. B. Uhrin, A remark to the paper of H. Hadwiger "Überdeckung des Raumes durch translationsgleiche Punktmengen und Nachbarnzahl", Monatsh. f. Mathematik, 104 (1987), 149-152.

8. H. Hadwiger, Überdeckung des Raumes durch translationsgleiche Punktmengen und Nachbarnzahl, Monatsh. f. Mathematik, 73 (1969), 213-217.

9. R.P. Bambah, A. Woods, H. Zassenhaus, Three proofs of Minkowski's second inequality in geometry of numbers, J. Austral. Math. Soc., 5 (1965), 453-462.

10. B. Uhrin, Some estimations useful in geometry of numbers, Period, Math. Hungar., 11 (1980), 95-103.

11. B. Uhrin, On a generalization of Minkowski convex body theorem, J. of Number Th., 13 (1981), 192-209.

DISPERSIONS, PROJECTIONS AND MEASURES OF COVERING IN EUCLIDEAN SPACES

B. Uhrin

Summary

Let $M \subset R^n$ be a closed (not necessarily discrete) subgroup of R^n (considering R^n as a locally compact Abelian group, where the Lebesgue measure μ is the Haar measure and the vector addition is the group operation). Let $A \subset R^m$ be a bounded L-measurable set and let $A+M$ denote the algebraic sum of A and M, i.e. the union of all sets of the form A+m, where m runs through M. The paper deals with the evaluation of $\mu((A+M) \cap G(r))$, where $G(r)$ is a ball of radius r (r is large comparing with the "size" of A). It is showed that $\mu((A+M) \cap G(r))$ is approximately equal to the product of two numbers, one of them depending on A and M only and the second one on r only. Denote the first one by $\alpha(M,A)$ (this is called the measure of covering). The paper is devoted to a detailed study of properties of $\alpha(M,A)$. Some identities (or equivalent forms) are proved for $\alpha(M,A)$. The main result is as follows: Let $K \subset R^n$ be a convex body. If the dimension of M is $n-k$, then $\alpha(M,tK)/t^k$ is an increasing function of t for $t \geq 0$. Many other results can be found in the paper, say, from the above property of $\alpha(M,tK)$ a sharpening of a classical theorem in the geometry of numbers (the successive minima theorem) easily follows.

DISZPERZIÓK, PROJEKCIÓK ÉS LEFEDÉSI MÉRTÉKEK EUKLIDESZI TEREKBEN

Uhrin Béla

Összefoglaló

Legyen $M \subset R^n$ egy zárt /nem feltétlenül diszkrét/ részcsoportja R^n-nek. Itt az R^n-et egy lokálisan kompakt Ábel csoportnak fogjuk fel, ahol a μ Lebesgue mérték a Haar-mérték és a vektorok összege a csoport müvelet. Legyen $A \subset R^n$ egy korlátos L-mérhető halmaz és jelöljük A+M-el az A és M algebrai összegét /azaz az A+m halmazok unióját, ahol m befutja az M-et/. A cikk azzal a kérdéssel foglalkozik, vajon "mennyire" fedi le az A+M halmaz az R^n-et, pontosabban meghatározandó a $\mu((A+M)\cap G(r))$ mérték, ahol $G(r)$ az r sugaru gömb /r nagy az A nagyságához képest/. A cikkben megmutatjuk, hogy a $\mu(A+M)\cap G(r))$ közelitőleg két szám szorzataként állitható elő, ahol az egyik csak az A-tól és M-től függ, jelöljük ezt $\alpha(M,A)$-val és a másik csak az r-től, jelöljük ezt $C(r)$-el. A cikk az $\alpha(M,A)$-nak /amelyet a lefedés mértékének nevezünk/ tulajdonságait vizsgálja. Az $\alpha(M,A)$-ra több azonosság van bizonyitva. Az egyik fő eredmény a következő: Legyen $K \subset R^n$ egy konvex test. Ha az M dimenziója n-k, akkor $\alpha(M,tK)/t^k$ monoton növekvő függvénye a t-nek, $t \geq 0$-ra. A cikkben sok egyéb eredmény is van, pl. az $\alpha(M,tK)$ fenti tulajdonságából a geometriai számelmélet egyik klasszikus tételének /a szukcessziv minimum tétel/ egy élesitése és kiterjesztése könnyen adódik.

MTA SZTAKI Közlemények 39/1988 p. 229-257.

ON THE STAR NUMBER OF A SET-LATTICE

B. UHRIN

Computer and Automation Institute,
Hungarian Academy of Sciences
1502 Budapest, Pf. 63., Hungary

1. Introduction

Let $\Lambda \subset R^n$ be the integer combination of n linearly independent vectors $b_1, b_2, \ldots, b_n \in R^n$ (the point lattice of full dimension). Let us call the collection $\mathcal{S}$ of the sets $\{S+u\}$, $u \in \Lambda$, the set-lattice, where $S \subset R^n$ is a bounded set ("Figurengitter" by Hadwiger [1]).

Denote by $T(\mathcal{S})$ the number of points $u \in \Lambda$ (the zero point Θ included) such that $S \cap (S+u) \neq \emptyset$. This number is called the star number of $\mathcal{S}$ ("Treffenzahl" by Hadwiger [2]). The number $T(\mathcal{S})$ has been introduced and studied when $\mathcal{S}$ is either a packing or covering of R^n (see, e.g. [3], [4] for more details), but it is obviously meaningful in the above most general situation as well.

The set S is called symmetric with centre $x \in R^n$ if $(S-x) = -(S-x)$.

If the centre of symmetric is the point Θ then S is called symmetric.

Erdős and Rogers [5] proved that if $S \subset R^n$ is a symmetric convex body (i.e. $S=-S$) such that $\mathcal{S}$ is a covering of R^n (i.e. $\bigcup_{u\in\Lambda}(S+u)=R^n$), then

$$(1.1)\qquad T(\mathcal{S}) \geqslant 2^{n+1}-1.$$

Groemer [6] extended (1.1) to any bounded $S\subset R^n$, i.e. he proved that (1.1) holds for any bounded symmetric S such that $\mathcal{S}$ covers the R^n. Groemer derived his result from a more general inequality. He first introduced the so called reduced star number $t(\mathcal{S})$ ("reduzierte Treffenzahl") as follows.

Given a $u\in\Lambda$ such that $S\cap(S+u)\neq\emptyset$, we take all $v\in\Lambda$ for which there is $w\in\Lambda$ such that $v+w=u$ and $(S+v)\cap(S+w)\neq\emptyset$. Denote by $d(u)$ the number of such v-s (this has been called by Groemer the degree of u).

If we collect all $u\in\Lambda$ such that $d(u)=k$, $k\geqslant 1$, then we get a decomposition of the set $\{u\in\Lambda:\ S\cap(S+u)\neq\emptyset\}$. Hence, denoting by $N_k(\mathcal{S})$ the number of u-s with $d(u)=k$, we get

$$(1.2)\qquad T(\mathcal{S}) = \sum_{k\geqslant 1} N_k(\mathcal{S}).$$

Then, by definition, the number

$$(1.3)\qquad t(\mathcal{S}):=\sum_{k\geqslant 1} k^{-1}\, N_k(\mathcal{S})$$

is called the reduced star number, [6].
It is clear that $d(u)\geqslant 2$ if $u\neq\theta$, $N_1(\mathcal{S})=0$ if $d(\theta)\geqslant 2$ and $N_1(\mathcal{S})=1$ if $d(\theta)=1$, consenquently

(1.4) $N_1(\mathcal{S}) \leq 1.$

Now, clearly ([6], p.23)

(1.5) $t(\mathcal{S}) = \frac{1}{2} N_1(\mathcal{S}) + \frac{1}{2}(N_1(\mathcal{S}) + N_2(\mathcal{S}) + \frac{2}{3} N_3(\mathcal{S}) + \cdots) \leq \frac{1}{2} N_1(\mathcal{S}) + \frac{1}{2} T(\mathcal{S})$,

hence (1.4) implies

(1.6) $T(\mathcal{S}) \geqslant 2t(\mathcal{S}) - 1$

The relations (1.5) and (1.4) show that equality is in (1.6) if and only if

(1.7) $N_1(\mathcal{S}) = 1$ and $N_k(\mathcal{S}) = 0$ for all $k > 2$.

The $\mathcal{S}$ fulfilling the conditions (1.7) has been called normal set-lattice, [6].

Groemer proved, using an identity for $t(\mathcal{S})$ ([6], Theorem 1) that if

(1.8) $S = -S$ and $\mathcal{S}$ covers the R^n,

then $t(\mathcal{S}) = 2^n$, yielding and extension of (1.1) to non-convex S.

The aim of this paper is to give refinements of the inequality (1.6) in the sense that, say,

(1.9) $T(\mathcal{S}) \geqslant M_1(\mathcal{S}) \geqslant M_2(\mathcal{S}) \geqslant 2t(\mathcal{S}) - 1$,

where $M_1(\mathcal{S})$, $M_2(\mathcal{S})$ are other well defined characteristics of $\mathcal{S}$.

Using (1.9) we get easily extensions and refinements of (1.1). Our inequalities can be used successfully also for giving new characterizations of normal set-lattices (as equality cases of (1.6)).

The clues to our results are two new type identities for $T(\mathcal{S})$. The first of them can be found also in [7], where it served as a tool for sharpening some upper estimations for $T(\mathcal{S})$ (see Section 6 for more details). We shall give also an identity for $N_k(\mathcal{S})$ that gives a new interesting insight into the Groemer's decomposition (1.2) and the quantity $t(\mathcal{S})$. Finally, our representation for $T(\mathcal{S})$ shows an interesting connection between lower estimations for $T(\mathcal{S})$ and a new sharper form of the classical Minkowski-Blichfeldt theorem proved in [8] (see Section 6).

2. The basic identities

Let $\mathcal{S} = \{(S+u) : u \in \Lambda\}$ be a set lattice in R^n where $S \subset R^n$ is a bounded set and $\Lambda \subset R^n$ a point-lattice generated by the basis $b_1, \dots, b_n \in R^n$. Denote $P := \{x \in R^n : x = \sum_{i=1}^{n} \lambda_i b_i, 0 \leq \lambda_i < 1, i=1,\dots,n\}$ (a unit cell of Λ). Denote by Λ' the lattice $\frac{1}{2}\Lambda$ i.e. Λ' is generated by the basis $b_1', b_2', \dots, b_n' \in R^n$, where $b_i' = 1/2\, b_i$, $i=1,2,\dots,n$. It is clear that $\Lambda \subset \Lambda'$.

One can see easily that the set $P' := P \cap \Lambda'$ is in a one-to-one correspondence with the quotient space Λ'/Λ (the set of different cosets $(\Lambda+x)$, $x \in \Lambda'$), i.e.

(2.1) $$\Lambda' = P' + \Lambda = \bigcup_{x \in P'} (\Lambda + x),$$

where the cosets $(\Lambda+x)$ are mutually disjoint for $x \in P'$. The canonical map $\psi: \Lambda' \to P' \sim \Lambda'/\Lambda$ is defined as

$$(2.2) \qquad \psi(u) = x, \text{ where } u \in \Lambda + x.$$

For any set $A \subset \Lambda'$ by definition

$$(2.3) \qquad \psi(A) := \bigcup_{a \in A} \psi(a).$$

The $\psi(A)$ is the canonical projection of A into $P' \sim \Lambda'/\Lambda$. One can see easily that

$$(2.4) \qquad \psi(A) = \{x \in P' : A \cap (\Lambda + x) \neq \emptyset\} = \bigcup_{u \in \Lambda} (A+u) \cap P'.$$

For any two sets $A, B \subseteq R^n$, $A+B$ means the algebraic (Minkowski) sum of the sets, i.e. the collection of points $a+b$, $a \in A$, $b \in B$. In particular $A-B := A+(-B)$. Our first identity is a straight consequence of the simple fact

$$(2.5) \qquad \{u \in \Lambda : S \cap (S+u) \neq \emptyset\} = (S-S) \cap \Lambda.$$

Hence

$$(2.6) \qquad T(\mathcal{S}) = |(S-S) \cap \Lambda|,$$

where $|A|$ denotes the cardinality of the finite set A.

Surprisingly enough, to our best knowledge, this almost trivial identity has not been used yet for the calculations concerning $T(\mathcal{S})$ (this identity has been successfully used also in [7]).

The second identity is formulated in the following

Theorem 2.1. For any bounded $S\subset R^n$ and any point lattice $\Lambda\subset R^n$,

$$(2.7)\qquad T(\mathcal{S}) = 1 + \sum_{i\geq 1}^{q(\mathcal{S})} \sum_{x\in\psi_i(\mathcal{S})} |2^{-i}(S-S)\cap(\Lambda+x)|,$$

where

$$(2.8)\qquad \psi_i(\mathcal{S}) := \psi\,(2^{-i}(S-S)\cap\Lambda^i)\setminus\{\Theta\},\quad i=1,2,\cdots,$$

$$(2.9)\qquad q\,(\mathcal{S}) := \min\{i:\ i\geq 0,\ 2^{-i}(S-S)\cap\Lambda = \{\Theta\}\}.$$

If $q(\mathcal{S})=0$ or $\psi_1(\mathcal{S})=\emptyset$ then the respective sums are considered by definition as zeros. □

Proof: Use (2.6). The condition $q(\mathcal{S})=0$ means $(S-S)\cap\Lambda=\{\Theta\}$, hence (2.7) is true. Let $q(\mathcal{S})>0$. It is clear that

$$(2.10)\qquad |(S-S)\cap\Lambda| = |2^{-1}(S-S)\cap\Lambda'|.$$

Using (2.1) we get

$$(2.11)\qquad 2^{-1}(S-S)\cap\Lambda' = 2^{-1}(S-S)\cap\Lambda \cup \bigcup_{\substack{x\in P'\\ x\neq\Theta}} (2^{-1}(S-S)\cap(\Lambda+x)),$$

where the non-empty sets in the union are mutually disjoint. By the relation (2.4)

$$(2.12)\qquad \psi(2^{-1}(S-S)\cap\Lambda') = \{x\in P':\ 2^{-1}(S-S)\cap\Lambda'\cap(\Lambda+x)\neq\emptyset\}.$$

But $\Lambda + x \subset \Lambda'$ for $x \in P'$, hence

$$(2.13) \qquad \psi(2^{-1}(S-S) \cap \Lambda') = \{x \in P' : 2^{-1}(S-S) \cap (\Lambda + x) \neq \emptyset\}.$$

The above identities yield

$$(2.14) \qquad T(\mathcal{S}) = \begin{cases} T(\mathcal{S}^{(1)}) + \sum_{x \in \psi_1(\mathcal{S})} |2^{-1}(S-S) \cap (\Lambda + x)| & \text{if } \psi_1(\mathcal{S}) \neq \emptyset, \\ T(\mathcal{S}^{(1)}) & \text{if } \psi_1(\mathcal{S}) = \emptyset, \end{cases}$$

where $\mathcal{S}^{(1)}$ is the set-lattice $\{(2^{-1}S+u) : u \in \Lambda\}$.
Denote by $\mathcal{S}^{(i)}$ the set-lattice $\{(2^{-i}S+u) : u \in \Lambda\}$.
Puting into (2.14) $\mathcal{S}^{(i)}$ instead of $\mathcal{S}$ and $\mathcal{S}^{(i+1)}$ instead of $S^{(1)}$ we get for all $i \geq 0$

$$(2.15) \qquad T(\mathcal{S}^{(i)}) = \begin{cases} T(\mathcal{S}^{(i+1)}) + \sum_{x \in \psi_{i+1}(\mathcal{S})} |2^{-(i+1)}(S-S) \cap (\Lambda + x)| & \text{if } \psi_{i+1}(\mathcal{S}) \neq \emptyset, \\ T(\mathcal{S}^{(i+1)}) & \text{if } \psi_{i+1}(\mathcal{S}) = \emptyset. \end{cases}$$

The later identity shows that

$$(2.16) \qquad 2^{-i}(S-S) \cap \Lambda = \{\Theta\} \Rightarrow 2^{-j}(S-S) \cap \Lambda = \{\Theta\} \quad \text{for all } j > i$$

and

$$(2.17) \qquad 2^{-i}(S-S) \cap \Lambda = \{\Theta\} \Rightarrow \psi_j(\mathcal{S}) = \emptyset \quad \text{for all } j > i.$$

These implications and the definition of $q(\mathcal{S})$ show that $T(\mathcal{S}^{(q(\mathcal{S}))}) = 1$ and $\psi_j(\mathcal{S}) = \emptyset$ for all $j > q(\mathcal{S})$.
Applying (2.15) successively we get (2.7). ∎