

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

KÖZLEMÉNYEK

1982. december

Szerkesztőbizottság:

GERTLER JÁNOS (felelős szerkesztő)
DEMETROVICS JÁNOS (titkár)
ARATO MÁTYÁS, BACH IVÁN, GEHÉR ISTVÁN,
GERGELY JOZSEF, KERESZTÉLY SÁNDOR, KNUTH ELŐD,
KRÁMLI ANDRÁS, PRÉKOPA ANDRÁS

Felélős kiadó:

VAMOS TIBOR

ISBN 963 311 146 3

ISSN 0133-7459

Főprint nyomda 83004

MAGYAN
TUDOMÁNYOS AKADEMIA
KÖNYVTÁRA

TARTALOMJEGYZÉK

J. Bak: Statistical and pattern recognition program packages for medical diagnosis and prognosis at the Semmelweis University of Medicine	13
Kismarty-Lechner I., Viszt É.: Biológiai analóg jelfeldolgozó laboratórium TPA/1 kisszámítógépre a SOTE számítástechnikai csoportban	19
Korányi L., M. Kopp: Vegetatív jelek analiziséről	27
Lengyel T.: A kanonikus korrelációanalízis alkalmazása szívkoszorúér megbetegedések előrejelzésében	35
Ratkó I.: Az 1977 óta működő számítógépes infarktus regiszter működése, annak főbb részei	45
Ratkó I., Ruda M.: Az MTA SZTAKI számítógépes statisztikai osztályán folyó orvosbiológiai kutatások rövid áttekintése	59
Simon P.: Az egészségügyi szolgálat gyógyító-megelőző tevékenysége információs rendszerének korszerűsítési problémái	63
Srajber B.: Az Egészségügyi minisztérium részére fejlesztett számítógéppel segített rendszerek	75
Beck Gy.: Toeplitz-típusú szalagmátrixszal leírható lineáris egyenletrendszerek megoldása a gyors Fourier-transzformáció segítségével	79

СОДЕРЖАНИЕ

Бак Й: Пакеты статистических программ и программ распознавания образов для медицинской диагностики и прогноза в Медицинском университете им. И. Семмельвейса	13
Кишмарти-Лехнер И., Вист Э.: Центральная лаборатория обработки аналоговых биосигналов Медицинского университета им. И. Семмельвейса	19
Корани Л., Копп М: Методы для анализа вегетативных сигналов	27
Леньдел Т.: Применение канонического корреляционного анализа	35
Ратко И.: Учёт инфарктов	45
Ратко И., Руда М.: Несколько замечаний о работе отдела статистики Исследовательского института вычислительной техники и автоматизации ВАН, связанных с биомедицинскими вопросами	59
Шимон П.: Вопросы системы информации лечебно-профилактической работы медицинской службы	63
Шрайбер Б.: Информационные системы для Министерства здравоохранения	75
Бецк Дь.: Эффективное решение ленточной системы управлений Теплица с помощью быстрого преобразования Фурье ..	79

C O N T E N T

P a g e

<i>J. Bak:</i> Statistical and pattern recognition program packages for medical diagnosis and prognosis at the Semmelweis University of Medicine	13
<i>I. Kismarty-Lechner, É. Viszt:</i> The central computerized laboratory system for the analysis of analog biological signals in the Semmelweis Medical University	19
<i>L. Korányi, M. Kopp:</i> Digital processing methods for analyzing vegetative signals	27
<i>T. Lengyel:</i> The application of canonical correlation analysis to predict coronaria thrombosis	35
<i>I. Ratkó:</i> Infarction registering system	45
<i>I. Ratkó, M. Ruda:</i> A short survey of the biomedical researches of the Computer and Automation Institute Hungarian Academy of Sciences	59
<i>P. Simon:</i> On how to make more up to date the information system of medical and preventive care in health service	63
<i>B. Srajber:</i> Some computer-aided system for the ministry of health	75
<i>G. Beck:</i> An efficient solution of banded Toeplitz systems by application of fast Fourier transform	79

T Á J É K O Z T A T Ó

A KNVVT ORVOSBIOLÓGIAI ADATOK SZÁMÍTÓGÉPES FELDOLGOZÁSÁVAL FOGLALKOZÓ MUNKACSOPORTJÁRÓL

A KNVVT /a szocialista országok tudományos akadémiáinak "A számítástechnika tudományos kérdései" bizottsága/ 1980-ban hozta létre 16. /"Orvosbiológiai adatok számítógépes feldolgozása" elnevezésű/ munkacsoportját. A csoport tagjai: Bulgária, Csehszlovákia, Lengyelország, Magyarország, a Szovjetunió és Kuba.

A munkacsoport feladata a nemzetközi együttműködés előmozdítása. Ennek érdekében meghatározta a több országot is érintő főbb kutatási témákat, a témát irányító és a résztvevő országokat. Természetesen bármelyik ország javasolhat újabb témákat vagy bekapcsolódhat a már meglévőkbe. A jelenlegi témák:

1. ORVOSBIOLÓGIAI ADATOK SZTATISZTIKAI FELDOLGOZÁSA

Ezt a témát Csehszlovákia és Magyarország irányítja, feladataink:

- a./ a meglévő statisztikai programok jegyzékének elkészítése,
- b./ a programok rövid jellemzése, felhasználási területük ismertetése,
- c./ konzultációk biztosítása a tagországok részére.

Ez utóbbihoz kapcsolódva 1981-re nyári iskolát tervezünk Csehszlovákiában.

2. BIOELEKTROMOS JELEK FELDOLGOZÁSA

A munkacsoport Bulgáriában hoz létre bázist ennek a kérdésnek a tanulmányozására. Bolgár irányítással Lengyelország és a Szovjetunió vesz részt egy elektromiográfiai jeleket feldolgozó programcsomag elkészítésében.

3. SZÁMITÓGÉP ALKALMAZÁSA A DIAGNÓZIS ÉS PROGNÓZIS TERÜLETÉN

Minden résztvevő országban működnek különböző diagnosztikai rendszerek, a további fejlesztés irányait a következőképpen tervezi a munkacsoport:

- a./ statisztikai módszereken alapuló diagnosztikai és prognosztikai rendszerek kifejlesztése kisméretű gépekre, magyar irányítással Bulgária, Csehszlovákia és Szovjetunió részvételével.
- b./ algebrai módszereken alapuló eljárások kifejlesztése /Lengyelország és Szovjetunió/,
- c./ diagnosztikai és prognosztikai rendszer létrehozása ritka betegségek esetére /melyekre a statisztikai módszerek nem alkalmazhatók/ Lengyelország és Szovjetunió részvételével,
- d./ lingvisztikán alapuló diagnosztika és prognosztika /Csehszlovákia/.

4. ORVOSBIOLÓGIAI FOLYAMATOK MATEMATIKAI MODELLEZÉSE

A következő modellek /és a hozzájuk tartozó programok/ létrehozását határozta el a munkacsoport:

- a./ májsejtek megbetegségéde /Lengyelország, Szovjetunió/,
- b./ a vázizomzat, aktivizálásakor fellépő ionos folyamatok /bolgár vezetéssel valamennyi ország/,

- c./ simaizomszövet sejtjeinek potenciáljai /Bulgária, Szovjetunió/,
- d./ a simaizomszövet összehúzódása elektromos stimuláció hatására /Bulgária/.

5. ORVOSI INFORMÁCIÓS RENDSZEREK

Lengyelországban kidolgozták a kórházi számítógépes rendszerek részletes terveit. Már működik a felvételi, a laboratóriumi és a patomorfológiai rendszer. Ezek interaktív módon működnek, latin nyelven a SNOMED szabványos orvosi nomenklatura alapján.

- a./ Lengyelországban bázist hoznak létre a munkacsoportban résztvevő országok tájékoztatására a klinikai és laboratóriumi orvosi információkat feldolgozó rendszerekről,
- b./ a patomorfológiai szabványos nyelv kidolgozása és bevezetése /lengyel irányítással valamennyi ország/,
- c./ az orvosbiológiai információfeldolgozás területén dolgozó szakemberek egységes nyilvántartásának létrehozása /Csehszlovákia/,
- d./ tipikus kórházi orvosi, információs rendszerrel kapcsolatos ajánlások kidolgozása /az 1983-ban Lengyelországban tartandó szimpóziumokon/.

6. ORVOSI INFORMÁCIÓK DIAGNÓZIS ELŐTTI ALANIZISE

A biztonságos diagnosztikának elengedhetetlen feltétele az adatok megfelelő ellenőrzése. Az erre szolgáló módszerek kidolgozásával Lengyelországban és a Szovjetunióban foglalkoznak. A rendszer élőszóval való működtetéséhez szükséges blokkot a Szovjetunió dolgozza ki.

7. SZOLGÁLTATÓ PROGRAMOK LÉTREHOZÁSA

- a./ orvosbiológiai /statisztikai/ adatok tárolása, szintaktikus és szemantikus ellenőrzése,
- b./ az adatok transzformációja, kiválasztása és csoportosítása további feldolgozások céljára,
- c./ alapvető statisztikai számítások elvégzése,
- d./ táblázatok készítése,
- e./ információs és nyilvántartó rendszerek létrehozása.

Magyar irányítással a csoport calamennyi tajának részvételével.

Kérjük azokat az intézményeket, amelyek valamilyen formában bekapcsolódnának a *hazai* munkacsoport tevékenységébe, vegyék fel Intézetünk - MTA SZTAKI - /Kerékfy Pál, Ratkó Istávn/ a kapcsolatot.

Jelen tanulmány az eddigi üléseken elhangzott, nyomtatásban még nem megjelent előadásokat tartalmazza.

On the Working Group 16 "Biomedical Information Processing"
of the KNVVT of the Academies of Sciences in Socialist Countries

*

О рабочей группе 16 "Обработка биомедицинских данных с помощью ЭВМ" Комиссии "Научные вопросы вычислительной техники" Академии наук социалистических стран.

Во время будапештского заседания РГ-16 "Обработка биомедицинских данных с помощью ЭВМ" Комиссии "Научные вопросы вычислительной техники" Академии наук социалистических стран были заслушаны доклады венгерских исследователей. Эти очерки содержат эти доклады.

STATISTICAL AND PATTERN RECOGNITION PROGRAM PACKAGES
FOR MEDICAL DIAGNOSIS AND PROGNOSIS AT THE SEMMELWEIS
UNIVERSITY OF MEDICINE

J. BAK

Semmelweis University of Medicine,
Computing Group

Medical diagnosis and prognosis based on laboratory tests and physical measurement are one of the most intriguing task of the biomedical application of pattern recognition techniques, discriminant and cluster analysis methods. These methods are in the field of mathematical statistics and a lot of statistical program packages have some classification algorithms themselves.

At Semmelweis University of Medicine we adapted a general purpose statistical program package, BMDP [1] /Biomedical Computer Programs, 1977/, for an ESZR-1020 computer. This program package contains statistical programs in the first place, however, among multivariate methods there are discriminant analysis and cluster programs for clustering cases and variables too. Among the programs there are some methods for data pre-processing, data reduction or transformation. These methods may be very important from the point of view of feature extraction. /It is to be noted that the methods the features are obtained by, are often intuitive and empirical, and their efficiency depends on the designer's knowledge and experience in the problem./

Besides BMDP we often apply SSP [2] /Scientific Sub-routine Package, IBM/. It contains about 250 subroutines, including factor analysis and discriminant analysis programs which are essential for pattern recognition.

These program packages usually have no supervised learning algorithms and these are only a few programs for unsupervised learning methods too.

In mathematical pattern recognition we want to get a decision rule which can classify examples of patterns quickly. A pattern recognition problem thus begins with class definition and labelling samples of the classes in some workable representations. The problem is solved when a decision rule has been derived to assign a unique label to new patterns.

For the statistical pattern recognition techniques we developed a program package, GALAXY [3]. It contains linear discriminant function, committee-machine algorithm /piecewise linear discriminant function/, first k-nearest neighbour decision rule, and potential function methods. The goals of such techniques are to identify the parameters /if any/ which can qualitatively distinguish the known groups and to select, if possible, a classification rule for identifying the known groups.

Besides learning algorithms, GALAXY contains some basic statistical programs for computing mean value, standard deviation, kurtosis, skewness, correlation coefficients, regression and empirical distribution. We don't require a rigorous data structure, i.e. data may be given on cards, magnetic tape or disc and there are no restrictions for the format of the data either.

This program can be operated with control cards.

GALAXY has been effectively applied for medical diagnosis and prognosis tasks, such as the prognosis for myocardial infarction patients [8], classification of sleep stages [9,10], and studying of acute cerebrovascular diseases.

We modified and adapted a nonhierarchical cluster program, DIDAY [4] developed at SZÁMKI /Budapest/ for an ESZR-1020 computer.

Most analysis techniques assume the homogeneity of the variables, whereas real data sets often have mixed variables. Our methods are applicable for continuous variables only. We have no methods for discrete, binary or dichotomous variables, although cases are frequently described by such variables.

/For instance, classifying the cases on the basis of medical records or giving parameters by coded data./

That is why we would like to complete our program packages with discrete discriminant analysis methods /W. Goldstein, [5]/ for solving problems of discrimination between groups with discrete multivariate observations, and with some hierarchical clustering algorithms for binary variables /SYN-TAX, [7]/. We have some experience with the medical applications of the program package ARTHUR [6] too, what was developed originally for chemical applications by L. Kowalsky.

REFERENCES

- [1] Dixon, W.J.; Brown, M.B.: Biomedical Computer Programs, P-Series, University of California Press, 1977, Berkeley.
- [2] System/360 Scientific Subroutine Package, Version III, Programmer's Manual /IBM/, 1968.
- [3] Bak, J.; Szadeczky-Kardoss, G.: Program Package for Classification Algorithms. Dokumentation /GALAXY/. March, 1977, SOTE.
- [4] Diday, E.; Schroeder, A.: The Dinamic Cluster Method in Pattern Recognition, Information Processing 74.
- [5] Goldstein, M.; Dillin, W.R.: Discrete Discriminant Analysis. John Wiley and Sons, New York, 1978.
- [6] Harper, A.M.; Duewer, D.L.; Kowalski, B.R.: ARTHUR, and Experimental Data Analysis: The Heuristic Use of a Polyalgorithm. Chemometrics: Theory and Application, 1970.

- [7] Podani, J.: Computer Program Package for Cluster Analysis in Ecology, Phytosociology and Taxonomy. Abstracta Botanica, Tomus VI., 1980, Budapest.
- [8] Jánosi, A.; Bajkai, G.; Bak, J.: Prognosis of Patients with Myocardial Infarction Treated in an Intensive Coronary Care Unit. Cardiologica Hungarica, 1979.
- [9] Pál, I.; Bak, J.; Halász, P.; Rajna, P.; Kundra, O.: Wide-band Spectra of Different SWS2 Stages are Different. In: Sleep, 1978. L. Popovics /Ed./, S. Karger, Basel. 4th Europ. Congr. of Sleep Research.
- [10] Bak, J.; Pál, I.; Halász, P.; Rajna, P.: Cluster Analysis of Broad-Band EEG Spectra in Identical Sleep Stages. Magyar EEG Társ. XXIV. évi tud. ülése, Debrecen, 1980. /XXIV. Congress of the Hungarian EEG Society, 1980, Debrecen/

ÖSSZEFOGLALÁS

J. Bak

Az orvostudomány számos területén, akár a differenciál diagnózis témakörében, akár a diagnóziskészítés folyamatának vizsgálatában felmerülnek olyan osztályozási vagy csoportba sorolási problémák, melyek igénylik az alakfelismerési módszerek alkalmazását. A SOTE-n kidolgoztunk egy alakfelismerési programokat tartalmazó rendszert, mely elsősorban tanuló algoritmusokat tartalmaz. Cluster analízis témakörből mások által fejlesztett programokat vettünk át és alkalmaztunk gyakorlati feladatokra. Néhány évvel ezelőtt adaptált statisztikai programrendszereink, mint a SSP, vagy BMDP is tartalmaznak olyan többváltozós módszereket /hierarchikus cluster algoritmusok változókra és esetekre, diszkriminancia és faktoranalízis/, melyek ezirányu munkáinkat nagymértékben elősegítik. A módszerek gyakorlati alkalmazásait illetően a prognózis területén a myocardiális infarctus lefolyását vizsgáltuk a kórházi időtartam /28 nap/ alatt, a diagnosztika területén pedig az akut cerebrovascularis kórképeket tanulmányoztuk. Jelananalízis témakörben az alvásfázisok automatikus felismerése és az alvás folyamat alatti fáziskülönbségek kimutatása jelentette sikeres alkalmazását az osztályozási módszereknek.

ПАКЕТЫ СТАТИСТИЧЕСКИХ ПРОГРАММ И ПРОГРАММ РАСПОЗНАВАНИЯ
ОБРАЗОВ ДЛЯ МЕДИЦИНСКОЙ ДИАГНОСТИКИ И ПРОГНОЗА В
МЕДИЦИНСКОМ УНИВЕРСИТЕТЕ им. И. СЕММЕЛЬВЕЙСА

Содержание

Во многих областях медицины, в том числе в области дифференциальной диагностики или при исследовании диагностического процесса возникают проблемы классификации и группировки объектов, требующие применения методов распознавания образов. В Медицинском Университете им. И.Семмельвейса разработана система программ распознавания образов, содержащая в первую очередь программы, реализующие обучающиеся алгоритмы. В области кластер-анализа применили для решения практических задач программы, взятые у других в готовом виде. Многомерные методы (как иерархический кластер-анализ переменных и событий, дискриминантный, факторный анализ) содержатся также в пакетах статистических программ /SSP, VMDF/, адаптированных у нас в последние годы, что значительно облегчает нашу работу в этом направлении. На практике эти методы были применены в решении следующих задач: в области прогноза исследовали течение инфаркта миокарда во время пребывания больного в стационаре (28 суток), а в области диагностики исследовали различные формы острых церебро-васкулярных заболеваний. В области анализа биосигналов методы классификации успешно применились при автоматическом распознавании фаз сна, и при выявлении различий между фазами сна.

THE CENTRAL COMPUTERIZED LABORATORY SYSTEM FOR THE ANALYSIS OF ANALOG BIOLOGICAL SIGNALS IN THE SEMMELWEIS MEDICAL UNIVERSITY

Ildikó Kismarty-Lechner, Éva Viszt
Computing Group, Semmelweis Medical University
Budapest, Hungary

To aid theoretical and clinical research work of the University we developed a general purpose laboratory system to process analog biological signals. Here are some of the topics solved within the system:

- analysis of EEG signals during oral glycerin therapy,
- quantitative analysis of the compound action potentials of the sympathetic efferent nerves,
- analysis of the spontaneous and evoked action potentials,
- Tensilon-effect on the tone and electrical activity of the outer eye muscles,
- spectral analysis of the airflow and oesophageal pressure signals,
- effect of the placental transfusion on the adaptation of newborns,
- evaluation of the intrauterine pressure and fetal heart rate signals.

The hardware configuration of our system is shown on Figure 1. For digitizing, storing, displaying and analysing analog signals the software was generally developed in assembly language. This is the basis for the solution of a number of specific problems. Particular problem can be solved with some supplementary programs which can be written in high level languages of TPAi or other computers (e.g. R20).

For A/D conversion

graphic digitizer (RA-01)

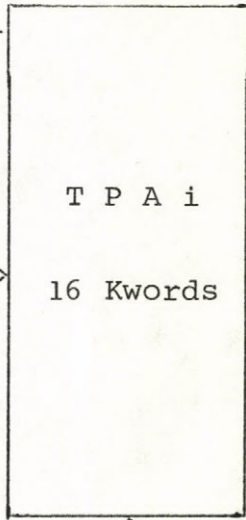
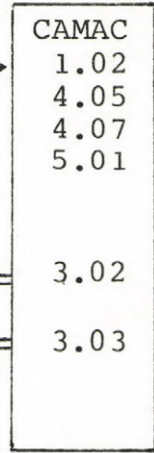
analog tape recorder

For D/A conversion

raster display (CAM 3.01)

X_Y recorder (BAK 5T)

polygraph



For data storage

magnetic tape units
(EC 5017)

fixed head disc
(DISCMOM 256K)

standard peripheral devices

Figure 1.
The hardware configuration of the computerized laboratory system

THE MAIN CHARACTERISTICS OF THE SYSTEM ARE:

- it is easy-to-learn; a conversational communication through a console display,
- it is easy-to-extend; modular construction (input, output, displaying and operating modules),
- serial measurement and evaluation,
- economic computation time,
- data from a maximum of 8 channels can be analysed simultaneously,
- special common data format in the core memory: single precision fixed point signed binary fraction with a variable scale factor (or exponent), the real and imaginary points having the same exponents,
- 4 to 4096 (by powers of 2) real or complex points can be processed at one (it is a block),
- data generally stroed on magnetic tapes (one volume comprises a number of files).

For digitizing the measured data the system has two analog input peripheral units:

- the CAMAC system is suitable for measuring voltages within the range of ± 5 V. The sampling frequencies can be deiscrete values provided by the CAMAC pulse generator. The maximal frequency that can be obtained in 10000/number of channels [Hz]. The sampled data are recorded on magnetic tapes.
- the RA-0= graphic digitizer can be used for digitizing data recorded on paper, e.g. with a compensograph.

Most often for the evaluation of the digitized signals the MATH program package written in assembly language is used. The most important operating subroutines (2) are the following: Fast Fourier Transform (FFT), inverse Fast Fourier Transform (IFFT), auto- and crosscorrelograms using FFT and IFFT. Generally the power-density spectra are calculated with the Fourier transformation method too. For FFT the Cooley-Tukey algorithm was applied (1). 1024 points may be transformed in 4.5 seconds, while 4096 points in about 10 seconds. The errors

caused by integral transformations are corrected with window-functions (by Bingham, Paresn, Hann, Hamming). There are some other operating subroutines e.g. for the selection of the evoked action potentials and other evoked signals. The MATH program package can be extended with new operating subroutines.

In addition, several programs in high level languages (TPAi BASIC, R20 FORTRAN) have been developed and/or adopted e.g.:

- Amplitude Analyser Program package (sampled mean, standard deviation, median, mean deviation, quantiles; empirical distribution function, central moment coefficients etc.),
- evaluation of the baseline and reference signals,
- calculation of the main points of positive and negative parts, latency time, zero-crossing, min. and max. values for the evoked action potentials,
- a special "peak-finding" program for locating fundamental and higher harmonics,
- the BMDP is the most important program package on R20,
- Time Series Analysis Program package on R20.

For displaying digitized signals three types of instruments are used:

- the CAMAC display: primarily for selecting and controlling. Signals from only 2 channels can be displayed at one time. Blocks are displayed one after the other. The user can stop, skip forward or backwards any of the displayed blocks.
- an X-Y plotter: mainly for documentary purpose. Its disadvantage is that it is relatively slow and the paper must be changed after each plotted figure.
- a polygraph: also for selecting and control purposes. Figures from a maximum of 4 channels can be displayed at once.

The connection towards the high level languages is achieved by the "pseudo-handlers" of the MATH program package. Pseudo-handlers have been developed for TPAi BASIC, R20 FORTRAN. This is very important because this keeps the system open. The input pseudo-handlers receive the different data formats and

convert them into core memory format which can be evaluated. The output pseudo-handlers perform the opposite function. The pseudo-handlers generally use the OS/i handlers but we had to develop some new ones for the non-standard peripherals e.g. RA-01. The pseudo-handlers are the input/output modules of the MATH program package therefore they are flexible and can be developed separately just as the operational and displaying modules.

SUMMARY

The central computerized laboratory system for the analysis of analog biological signals is based on a TPAi and a CAMAC hardware. The programs are generally written in assembly language. Due to the modular arrangement the system can easily be expanded. The OS/i operational system has been modified and expanded with new programs. Magnetic tapes provide the connection towards other computers.

In developing the system Lóránt Ormai, Éva Keszthelyi, Zoltán BYdeskuty and László Szekeres helped the authors.

REFERENCES

- [1] Robert K. Otnes, Loren Enochson: Digital time series analysis
- [2] Kismarty-Lechner Ildikó: MATH dokumentáció (manuscript)
- [3] Viszt Éva: EAN-DIG dokumentáció (manuscript)
- [4] Az analóg laboratórium felhasználói kézikönyve; 1979 (manuscript)

ЦЕНТРАЛЬНАЯ ЛАБОРАТОРИЯ ОБРАБОТКИ АНАЛОГОВЫХ БИОСИГНАЛОВ
МЕДИЦИНСКОГО УНИВЕРСИТЕТА ИМ. И. СЕММЕЛЬВЕЙСА

На базе ЭВМ типа ТРА/1 Группы Вычислительной Техники Мед. Унив. им. И. Семмельвейса с целью поддержки теоретической и клинической исследовательской работы университета создана лаборатория обработки аналоговых сигналов. Различные операции над сигналами производятся с помощью системы КАМАК, а для хранения информации установлены накопители на магнитных лентах. Ввод информации осуществляется с аналоговых магнитографов или с графического материала с помощью аналого-цифрового преобразователя РА-01. Устройствами вывода служат графопостроитель, дисплей системы КАМАК и 4-х-канальное самопишущее устройство.

Существующая в настоящее время система математического обеспечения, созданная с учетом самых различных требований пользователей имеет модулярную структуру, позволяет серийные измерения с последующей обработкой. Возможности:

- сбор данных с А/Ц преобразованием на нескольких каналах одновременно с различной частотой дискретизации;
- собственная система обработки /быстрое преобразование Фурье, обратное преобразование Фурье, автокоррелограммы, взаимные коррелограммы, различные спектральные весовые функции, и т.д./;
- вывод данных /т.е. построение кривых/ после А/Ц преобразования, а также результатов обработки;
- переход к программам, выполненным на алгоритмических языках высокого уровня /на ЭВМ ТРА/1 и ЕС-1020/.

Программы системы мат. обеспечения написаны на языке Ассембли, благодаря чему продолжительность их прогона весьма

удовлетворительна. Разделение операций ввода и вывода от операций собственно обработки позволяет дальнейшее развитие системы без особенных сложностей. Система доступна пользователям, не располагающим специальными знаниями и навыками в области вычислительной техники.

ÖSSZEFOGLALÁS

BIOLÓGIAI ANALÓG JELFELDOLGOZÓ LABORATÓRIUM TPA/I KISSZÁMÍTÓGÉPRE A SOTE SZÁMÍTÁSTECHNIKAI CSOPORTBAN

Kismarty-Lechner Ildikó, Viszt Éva

A SOTE Számítástechnikai Csoport TPA/i kisszámítógépére alapozva - az Egyetemen folyó elméleti és klinikai kutató munka segítésére - analóg jelfeldolgozó laboratóriumot hoztunk létre. A jelekkel kapcsolatos különböző feladatok elvégzésére CAMAC rendszert, adattárolásra mágnesszalagos egységet használunk. Adatainkat analóg magnetofonról és RA-01 rajzdigitalizálóval regisztrátumról visszük be. Megjelenítésre CAMAC display, X-Y rajzoló és négy csatornás gyorsregisztráló használható.

A sokrétű felhasználói igényeket figyelembe véve alakítottuk ki programrendszerünket, mely moduláris felépítésű, sorozatmérésre és értékelésre alkalmas. Szolgáltatásai:

- több csatornás mérés-adatgyűjtés változtatható mintavételi frekvenciával,
- saját értékelő rendszer /FFT, IFFT, autó- és keresztkorrelogram, ablakfüggvények, marker szerinti válogató stb./,
- digitalizált és értékelt jelek megjelenítése,
- csatlakozási lehetőség a TPA/i és R20 -s számítógépek magasszintű programnyelvei felé.

A softwaret assembler nyelven irtuk, így a gutási idők kedvezőek. Az input, output és műveleti részek szétválasztásával a rendszer könnyen továbbfejleszthető. Használata számítógépes ismeretekkel nem rendelkezők számára is elég könnyen elsajátítható.

DIGITAL PROCESSING METHODS FOR ANALYZING VEGETATIVE SIGNALS

L. Korányi
Medicor Works

M. Kopp
Semmelweis University of Medicine
Budapest Hungary

The following paper deals with some of the results of a joint research of Medicor Works and Semmelweis University of Medicine. Methods, similar to the described ones are now being used in recently developed micro-computer controlled medical instruments of Medicor Works.

The aim of our research was to identify the characteristics of two so called vegetative diseases: the hypertension and the ulcus duodeni; in their early stages, and to compare these characteristics between the two groups of patients and a control group. We even limited the scope of our investigations to the analysis of some significant response patterns evoked by psychophysiological stimuli.

ABOUT THE COURSE OF OUR MEASUREMENTS

In order to analyse the vegetative reactions, 3 signals were measured: The galvanic skin resistance (GSR), the pulse curve and electromyogram (EMG), all with surface electrodes. (fig.1.) The subjects were seated in a comfortable armchair in a sound protected, dark room. After 7 minutes of relaxation the effects of verbal, noise and light stimuli were measured. The stimuli followed each other with an interval of about 45 seconds, played from tape-recorder. The electrodes were fixed on the supported hand of the subject. The 3 signals had been

recorded on an analogue FM tape-recorder, together with a fourth one, which served as synchron channel during the processing phases. This way we gained about each subject approximately 4 times 1000 seconds long analogue recordings. The sampling and all of the processing was carried out with the aid of a PDP 11/45 installation (fig.2.) by programs written in BASIC-PLUS language.

With the sampling frequency of 21 c/s, we had 4 times 18 Kwords digital samples about each subject. The total amount of data, stored on magnetic tape - was about 20 Mbytes.

In the following sections of my paper, I shall deal with two methods of the analysis, which proved to be specially useful and yielded interesting results.

1. INTERACTIVE BAYES-DECISION

An interactive program was developed for the analysis of the GSR curves, using the Bayes decision.

Let us consider the binary decision problem briefly. (fig. 3.) The two density functions represent the two possible choices or hypotheses. The decision problem is equivalent to choosing one of the two density functions on the basis of one or more observations. It is obvious that decision problem arises only if the observed phenomenon is of probabilistic nature.

If we assume a threshold at ξ_T , then if the observation α is less than ξ_T , we accept H_0 , if α is greater than ξ_T , we accept H_1 . To determine the optimal threshold, we define the Bayes risk B , as the expected value of the cost for the two possible types of errors. The best threshold is the one, where the Bayes risk B , is minimal.

$$B = C_{10} \cdot P_{H_0} + \int_{-\infty}^{\xi_T} [-C_{10} \cdot P_{H_0} p_0(\alpha|H_0) + C_{01} \cdot P_{H_1} p_1(\alpha|H_1)] d\alpha \quad (1)$$

If we differentiate the expression of B with respect to T , we get that the necessary condition of the minimum is:

$$C_{01} \cdot P_{H_1} \cdot p_1(\alpha | H_1) - C_{10} P_{H_0} p_1(\alpha | H_0) = 0 \quad (2)$$

From (2) the Bayes-decision is:

$$\frac{p_1(\alpha | H_1)}{p_0(\alpha | H_0)} \begin{matrix} \text{accept } H_1 \\ \geq \\ \text{accept } H_0 \end{matrix} \frac{C_{10} P_{H_0}}{C_{01} P_{H_1}} \quad (3)$$

Where C_{01} and C_{10} are different costs associated with the errors.

During the analysis of the GSR signals we had to solve the problem, how to handle the great amount of samples and data. For this purpose we developed an interactive method, where the operator has to intervene only in the case when the computer decision has to be overridden. In the GSR signals we looked for the resistance-fall waves and for their characteristics. To be able to use the Bayes-decision efficiently, a transformation had to be carried out on the signal. (fig.4.) As it can be seen we produced such a point series which consists of only the differences of the local maximums and minimums of the original sample. We considered, that in order to determine the responses for a stimulus in the GSR curve, we have to determine in the compressed sample the places of great differences, and the small differences are considered as measurement noise from our point of view. This is a simple decision problem, where we want to detect a constant signal in the presence of additive noise. The distribution of the noise is assumed to be zero-mean gaussian, and the amplitude of the signal is m . In this case the two density functions belonging to the two hypotheses are:

$$p_0(\alpha | H_0) = \frac{1}{\sqrt{2\pi} \delta} e^{-\alpha^2 / 2\delta^2} \quad (4)$$

$$p_1(\alpha | H_1) = \frac{1}{\sqrt{2\pi} \delta} e^{-(\alpha - m)^2 / 2\delta^2} \quad (5)$$

And the Bayes-decision test in this case:

$$\alpha \begin{matrix} \text{accept } H_1 \\ > \\ \text{accept } H_0 \end{matrix} \frac{m}{2} + \frac{\delta^2}{m} \ln \frac{P_{H_0}^C}{P_{H_1}^C} \quad (6)$$

With the aid of the maximum-minimum compression, described above, we achieved that this decision criterion had to be used for a single sample only 2-3 hundred times instead of 18 000 times, if used it for each sample point.

About the operation of the program: The original curve was displayed on a raster-display and a sign was put on each point found, one after the other. The operator could change the place of the point with the aid of a moving cursor. After identifying the desired interval, the program calculated further characteristics of the reponse, like duration, slopes, integral value, reaction time etc.

Figure 5. shows a screen, where all answers were found automatically, indicated by the short vertical lines set on the curve. (The vertical lines below indicate the places of stimuli.) Figure 6. shows a screen where operator intervention was necessary (dotted vertical lines).

In about 20 to 25% of the cases was necessary to override the decision of the algorithm manually.

2. FREQUENCY DOMAIN DATA REDUCTION

In the case of the pulse curve analysis a completely different method was used, which is interesting from the data reduction point of view. Here I would like to emphasize that in this case our purpose was to compare the reactions between the groups, so we did not even strive for quantitative results.

The processing of the pulse curve was carried out in two steps. The task of the first step can be seen on fig. 7. The peak values and locations of the single pulse waves were determined by a peak detections algorithm. The outputs of this program are the peak values and the beat-to-beat pulse times.

This is presented on fig. 8. The size of these outputs are about 10% of the size of the input files.

To characterize the relation of the stimuli and the responses in the curves, we used the cross-correlation function. For this purpose we constructed a square-wave function from the synchronizing signal as it can be seen on fig. 9. Since the information about the responses in the cross-correlation function is still in a "distributed" manner, there was a need for further data compression to make possible the comparison of the great amount of samples. In order to do this we computed the Fourier-transform of the cross-correlation functions, that is the cross-power spectra. Figure 10. gives an example for these spectra. We found that the first 30 spectrum lines contain over 75 % of the total energy, these 30 data are sufficiently enough to characterize the vegetative response function in the frequency-domain. By sufficient we mean that it can be used for differentiating curves belonging to different groups. With the help of this method we could make the feature selection desired, with the 0.2% of the original 18 Kw of data. Figure 11. shows to typical spectra belonging to the hypertensive and to the control group.

To compare the spectra between groups we used the relative energy contents of the spectrum lines. Even with this relatively rough method there were significant differences found between the spectra of the two groupes, referred above, at 1/40 c/s which is the stimulation frequency, and at the 3. and 5. harmonics. There were no significant differences between the spectra of the control and the ulcer duodeni group.

Without explaining the medical significance of our findings, the two methods which were presented very briefly proved to be efficient digital analysis methods for these common physiological signals. Some of this methods are being, or can be applied in biomedical instruments as well.

REFERENCES

Bendat J.S., Piersol A.G.: Random data: Analysis and Measurement Procedures, Wiley - Interscience, New York, 1971.

Sage A.P., Melsa J.L.: Estimation Theory with Applications to Communications and Control, McGraw-Hill, New-York.

МЕТОДЫ ДЛЯ АНАЛИЗА ВЕГЕТАТИВНЫХ СИГНАЛОВ

В настоящей статье приведены некоторые результаты совместных исследований, проведенных Объединением "Медикор" и Университетом Медицинских Наук им. Семмельвейса. Аналогичные указанному методу анализа сигналов процедуры нашли применение и в Объединении "Медикор" в разработанных в последние годы микропроцессорных медицинских приборах.

ÖSSZEFOGLALÁS

VEGETATIV JELEK ANALIZISÉRŐL

Jelen cikk a Medicor Művek és a Semmelweis Orvostudományi Egyetem közös kutatásának néhány eredményét tárgyalja. Az ismertetett jelanalízis módszerekhez hasonló eljárások alkalmazásra kerültek a Medicor Műveknél az utóbbi időben kifejlesztett mikroszámítógépes orvosi műszerekben is.

THE APPLICATION OF CANONICAL CORRELATION ANALYSIS TO PREDICT CORONARIA THROMBOSIS

Tamás Lengyel

Computing and Automation Institute of
Hungarian Academy of Science
Budapest

1. INTRODUCTION

We should like to call your attention for an application of canonical correlation analysis which is useful to predict some aspects of complex phenomena.

We dealt with prediction of a special heart-disease: the so-called coronaria thrombosis.

The data were given by the National Institute of Cardiology (Budapest). Ten variables concerning the physical status, smoking customs were measured in 3 villages, at 3 times by 5 years, in a population with size of 1082 people ([3]).

Our hypothesis was that data have information not only about the actual status but about their possible changes expected in a short term. So, we were interested in studying connection between the consequent measures and we aimed at using these relations to predict measures and status by the same variables but which had been measured 5 years earlier.

One can solve problem like mentioned above by means of canonical correlation analysis.

This method seems us to be new in the field of analysis of very short time series.

In the first part of my paper I sketch out the concept of canonical correlation analysis then I present its application to our problem. Because of the shortage of

applicable data I had to restrict myself concerning the size of analysis.

2. THE CONCEPT OF CANONICAL CORRELATION ANALYSIS

Let us denote two multidimensional stochastic vector variables by \underline{U}_1 and \underline{U}_2 . We will study the relation(s) between \underline{U}_1 and \underline{U}_2 . We suppose that their components are standardized, i.e. the means and the variances equal to 0 and 1, respectively.

Let us suppose as well that \underline{U}_1 and \underline{U}_2 are in R^q and R^{p-q} , respectively ($p > q > 0$). Let us denote the common covariation matrix of \underline{U}_i and \underline{U}_j by $\Sigma_{ij} = \text{cov}(\underline{U}_i, \underline{U}_j)$ ($i, j = 1, 2$). In this case the covariation equals to the correlation.

We suppose that

$$\begin{aligned}\text{rank } (\Sigma_{11}) &= q, \\ \text{rank } (\Sigma_{22}) &= p - q,\end{aligned}$$

where $\text{rank } (A)$ means the rank of matrix A .

We introduce the following notations

$$\begin{aligned}m &:= \min \{q, p - q\} \\ k &:= \text{rank } (\Sigma_{12})\end{aligned}$$

A' means the transponent of A ,

A^{-1} means the inverse of a quadratic matrix A .

2.1 There is a well known measurement between 2 one dimensional stochastic variables, the so-called correlation coefficient.

We can write the (minimal least square) regression equation in the form

$$\begin{aligned}\hat{U}_2 &= rU_1 \\ E(U_2 - \hat{U}_2) &= 1 - r^2,\end{aligned}$$

where $r = \text{cov}(U_1, U_2)$.

We can interpret this fact as U_1 explains a quantity of r^2 from the variance of U_2 .

2.2 Furthermore, we know the multiple correlation coefficient as a good measure of relation between an one dimensional (U_2) and a multidimensional (\underline{U}_1) stoch.var. It equals to the correlation $r(U_2, \hat{U}_2)$ of U_2 and \hat{U}_2 where \hat{U}_2 denotes the best estimate of U_2 by means of linear regression with respect to \underline{U}_1 . However, this $r(U_2, \hat{U}_2)$ is equal to the maximum correlation between the linear functions of \underline{U}_1 and U_2 , and

$$E(U_2 - \hat{U}_2)^2 = 1 - R_{U_2 \cdot \underline{U}_1}^2,$$

where $R_{U_2 \cdot \underline{U}_1}$ is the multiple correlation coefficient of U_2 and \underline{U}_1 .

The above note on the explanation of variance holds.

2.3 Generally, we can describe the relation between two stochastic vector variables by the maximal correlation between their linear functions. This number is called canonical correlation coefficient. One can realize that this method is a natural generalization of the concept of correlation.

We can define $m := \min\{q, p - q\}$ canonical correlation coefficients and factors by analitical method.

We are able to pass over to a coordinate system (or factor space) where the components of \underline{U}_1 and \underline{U}_2 are uncorrelated except those coordinates which have the same indices and which have considerable correlations. In this space from the variance of the i^{th} component of \underline{U}_1 the \underline{U}_2 explains exactly the same quantity as its i^{th} coordanite does.

In analitical terms: the first canonical correlation equals to

$$\varphi_1 = \max r(\underline{L}'_1 \underline{U}_1, \underline{M}'_1 \underline{U}_2)$$

$$\underline{L}_1 \in R^q, \underline{M}_1 \in R^{p-q}$$

$$D^2(\underline{L}'_1 \underline{U}_1) = D^2(\underline{M}'_1 \underline{U}_2) = 1,$$

where $r(\xi, \eta)$ denotes the correlation between ξ and η , and $D^2(\xi)$ denotes the variance of ξ .

It is easy to prove that this minimum is reached. The coordinates of $\underline{L}'_1 \underline{U}_1$ are called first left-hand-side canonical factors and the coordinates of $\underline{M}'_1 \underline{U}_2$ are called first right-hand-side canonical factors.

We can get by the method of Lagrange multiplication ([8])

$$(\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \varphi_1^2 \Sigma_{22}) \underline{M}_i = 0 \text{ and}$$

$$(\Sigma_{12} \Sigma_{22} \Sigma_{21} - \varphi_1^2 \Sigma_{11}) \underline{L}_i = 0.$$

Let us introduce the following notations (where $\underline{L}'_i \underline{U}_1$ and $\underline{M}'_i \underline{U}_2$ are the proper side canonical factors ($i=1, 2, \dots, m$))

$$L = (\underline{L}_1 | \underline{L}_2 | \dots | \underline{L}_m)$$

$$M = (\underline{M}_1 | \underline{M}_2 | \dots | \underline{M}_m)$$

$$\Lambda = \text{diag}(\varphi_1, \varphi_2, \dots, \varphi_m) \quad (\varphi_i \geq 0)$$

If we suppose that $p=2q$ and $k=m(=q)$ then reading by rows

$$(1) \quad \widehat{M}' \underline{U}_2 = \Lambda L' \underline{U}_1$$

where $\widehat{M}' \underline{U}_2$ = the best least square linear regression with respect to $\{\underline{L}'_1 \underline{U}_1, \underline{L}'_2 \underline{U}_1, \dots, \underline{L}'_m \underline{U}_1\}$, and at the same time (by rows)

$$(2) \quad \widehat{\underline{U}}_2 = M'^{-1} \Lambda L' \underline{U}_1$$

where $\widehat{\underline{U}}_2$ = the best least square linear regression with respect to $\{\underline{L}'_1 \underline{U}_1, \underline{L}'_2 \underline{U}_1, \dots, \underline{L}'_m \underline{U}_1\}$.

We try to explain the relations between \underline{U}_1 and \underline{U}_2 by means of generalization of explanation of variance. We describe these relations by 2 numbers: each of them show the average (mean) quantity which we explained from the sum of variances of components of one vector variable by the other:

$$R_{\underline{U}_2 \cdot \underline{U}_1} \quad \text{and} \quad R_{\underline{U}_1 \cdot \underline{U}_2} .$$

These numbers generally are not equal and they depend on the coordinate system. We consider 2 systems:

a/ in these space of canonical factors

$$R_{M' \underline{U}_2 \cdot L' \underline{U}_1}, \quad R_{L' \underline{U}_1 \cdot M' \underline{U}_2} .$$

In this case we can write

$$\begin{aligned} \hat{M' \underline{U}_2} &= \Lambda L' \underline{U}_1, \\ \hat{L' \underline{U}_1} &= \Lambda M' \underline{U}_2 \quad \text{and} \end{aligned}$$

$$\varphi^2 = R_{M' \underline{U}_2 \cdot L' \underline{U}_1} = R_{L' \underline{U}_1 \cdot M' \underline{U}_2} = \frac{1}{k} \sum_{i=1}^k \varphi_i^2 .$$

b/ \underline{U}_2 in the original space and \underline{U}_1 in the space of canonical factors, and similarly \underline{U}_1 in the original space and \underline{U}_2 in the space of canonical factors

$$\begin{aligned} \hat{\underline{U}_2} &= M'^{-1} \Lambda L' \underline{U}_1, \\ \hat{\underline{U}_1} &= L'^{-1} \Lambda M' \underline{U}_2. \end{aligned}$$

In this case

$$R_{\underline{U}_2 \cdot \underline{U}_1} \neq R_{\underline{U}_1 \cdot \underline{U}_2}$$

We have to note that there is a clear connection between factor and canonical correlation analysis. Very similar optimizing features hold for both of them ([8], [4]).

There is a more complex explanation of canonical correlations the so-called redundancy analysis ([2]).

3. OUR RESULTS

Let us denote the measurement vector by \underline{U}_1 , and the vector for the same person measured five years later by \underline{U}_2 .

Our aim is clear:

we have to analyze the connection between \underline{U}_1 and \underline{U}_2 , and construct an estimation of \underline{U}_2 by \underline{U}_1 . We chose the following 6 variables:

- Broca-index
- Systolic blood-pressure
- Diastolic blood-pressure
- Vitalcapacity
- Serum Cholesterol
- Body mass index.

We constructed subsets of the sample. We found 209 individuals to be healthy during the 10 year period. 28 patients had already been ill at the first time. 17 persons became ill when their data were measured at second time.

We had to control whether the healthy and sick persons could be separated by discriminantial analysis on the basis of this sample.

The first-kind and second-kind error were about 30-31 %.

We cut the persons for 3 non-disjoint groups:

- HHH - healthy
- HS - people who became ill during the medical examinations
- HHH&HS - all of the individuals

At the beginning \underline{U}_1 and 5 years later \underline{U}_2 measure the data. In all 3 cases we established the estimation of the $\hat{\underline{U}}_2$ by (2). Table 1 shows the suitable correlation-type measurement numbers and their dependencies on chosen coordinate system. We classified these $\hat{\underline{U}}_2$ by means of discriminant analysis. Our result shows that the third regression gives the "minimal error" (Table 2).

So, we proved that by this method we could analyze the relation between the actual and expected status, and could state the proper conclusions with the help of regressions. The surprise of the analysis was that this relation was higher for people which were going to be sick than for healthy people.

	$R_{\underline{U}_1 \cdot \underline{U}_2}$	$R_{\underline{U}_2 \cdot \underline{U}_1}$
HHH	0.574	0.570
HS	0.734	0.735
HHH&HS	0.579	0.573

Table 1.

	first-kind error	second-kind error
HHH	0.25	0.35
HS	0.93	0.06
HHH&HS	0.24	0.35

Table 2.

LITERATURE

- 1 ANDERBERG, M.R., Cluster Analysis for Applications (Academix Press, New York-London, 1973)
- 2 COOLEY, W.W. and LOHNES, P.R., Multivariate Data Analysis (John Wiley and Sons, New York, 1971).
- 3 LENGYEL, T., "A kanonikus korrelációanalízis alkalmazása szivkoszorúér - megbetegedések előrejelzésére", Számítástechnikai és kibernetikai módszerek alkalmazása az orvostudományban és a biológiában, 8. Neumann Kollokvium, Szeged, 1977, 11-17.
- 4 LENGYEL, T., "A kanonikus korrelációanalízis és néhány kapcsolódó probléma (The canonical correlation analysis and some related problems)", Alkalmazott Matematikai Lapok 5 (1979).
- 5 MILLER, J.K., "The development and application of bivariate correlation: a measure of statistical association between bivariate measurement sets", Ed.D. Dissertation, Faculty of Educational Studies, State University of New York at Buffalo, 1969.
- 6 STEWART, D.K. and LOVE, W.A., "A general canonical correlation index", Psychological Bulletin 70 (1968) 160-163.
- 7 TUSNÁDY, G., "Mátrixok szinguláris felbontása", Alkalmazott Matematikai Lapok 5 (1979).
- 8 RAO, C.R., Linear Statistical Inference and Its Applications (John Wiley and Sons, New York, 1965).

Lengyel Tamás

A kanonikus korrelációanalízis alkalmazása
szívkoszorúér megbetegedések előrejelzésében

Egy olyan módszer alkalmazására adunk példát, ami az ilyen jellegű vizsgálatok körében ujszerű. Rövid idősorok esetében várható, hogy ez a módszer a vizsgálati adatok analizálásának hatékony segédeszköze lesz. A cikkben ismertetjük a kanonikus analízis elvét, egy konkrét példán keresztül bemutatjuk alkalmazását.

Megjegyezzük, hogy kevés értékelhető adat állt rendelkezésünkre, s ezért bizonyos egyszerűsítésekre kényszerültünk. A módszer további finomítására is van lehetőség.

Тамаш Лендел

Применение канонического корреляционного анализа

Автор даёт новый метод для применения канонического корреляционного анализа.

УЧЁТ ИНФАРКТОВ

ИШТВАН РАТКО

1. ВВЕДЕНИЕ

В Венгрии с 1970 года создана Южно-Пештская служба учета инфарктов, которая обеспечивает ознакомление с изменениями инфарктной заболеваемости и создает основу для оценки эффективности систем медицинского обслуживания и программ здравоохранения.

В 1976 году, при поддержке Министерства здравоохранения, мы начали разработку системы управления базой данных для регистрации инфарктов, требующей довольно большую администрацию. Эта система дала возможность более точной и быстрой оценки данных.

Изучение частоты случаев инфаркта — единственная возможность определения причин изменений смертности. В некоторых странах смертность от сердечной недостаточности значительно снизилась. Причина изменений смертности принимается как результат изменений лечения и обслуживания; профилактические меры тоже оказывают влияние. Определение причины не академический вопрос, так как материальные средства здравоохранения даже в самых богатых странах ограниченные, применение самых эффективных проблематично на практике. Когда смертность снижается при постоянной частоте случаев, это снижение приписывается скорее улучшению обслуживания больных /интенсивные отделения, операции *koronaria bypass*, новые лекарства и др./ пока в случае преобладания снижения частоты заболеваний, уменьшение смертности происходит благодаря вышеупомянутому положительному влиянию профилактических мер.

Из-за ожидающихся в 80-х годах изменений частоты заболеваний Всесоюзная Организация здравоохранения /ВОЗ/ предлагает странам-членам внедрение мониторной системы.

Разработанная нами система фигурирует в документации ВОЗ, как одна из предложенных для обсуждения программ.

Статистика смертности как в Венгрии, так и в развитых странах однозначно показывает значение сердечно-сосудистых заболеваний, в том числе и острого инфаркта миокарда /ОИМ/. У нас причиной половины общей смертности являются сердечно-сосудистые заболевания, а процент умерших от острого инфаркта миокарда выше 10. Распространенность болезни и инвалидность в результате нее в молодом трудоспособном возрасте делает борьбу с ней еще более неотложной.

Первым шагом этого по предложению ВОЗ явилось создание Службы учета инфарктов на территории медицинского обслуживания Южного Пешта. С 1972 года продолжаем работу в упрощенной форме.

2. ЦЕЛЬ УЧЕТА ИНФАРКТОВ

1/ Определить частоту заболеваний ОИМ, морбитальность, летальность среди населения точно определенной территории.

2/ Изучить природу болезни, в особенности раннего, догоспитального ее периода, значение предшествующих приступа симптомов, профилактику предполагаемых факторов риска, клинические формы и патанатомию болезни.

3/ Поднять проблемы обслуживания случаев ОИМ; в ост-
ром периоде в больнице и за пределами ее, а также после вы-
писки из больницы.

4/ Проследить за изменениями природы, смертности, ле-
тальности, частоты болезни в течении годов, под влиянием
условий жизни и медицинского обслуживания.

5/ Создать такую модель обслуживания, которая будет
информационной основой для дальнейших исследований, оцен-
ки организационных программ здравоохранения и в качестве
оперативной единицы станет частью кардиологического обслу-
живания.

3. РАБОТА И ОРГАНИЗАЦИЯ СЛУЖБЫ УЧЕТА ИНФАРКТОВ /СУИ/

Служба учета инфарктов работает на территории Южно-
пештского медицинского обслуживания, которая охватывает
IX, X, XUll, XUlll, XIX и XX-ые районы. Население террито-
рии полмиллиона.

СУИ разыскивает и берет на учет лица с острым инфарк-
том миокарда или с подозрением на него и каждого человека,
внезапно и неожиданно скончавшегося.

Ценность данных зависит от полноты учета, как и в
любом эпидемиологическом анализе. Для обеспечения полноты,
чтобы каждый случай, подозрительный на инфаркт был учтен,
мы применяем следующие методы:

- 1/ Еще в подготовительном этапе, в результате двух- и многосторонних консультаций, договорились о соответствующем сотрудничестве с работающими на территории врачами участков и больниц. Подробным образом установили хорошую связь с руководителями и врачами относящихся к другим министерствам больниц, куда из-за некоторых привилегий больные территории могут попасть. Большую помощь нам оказали Исполком отдела здравоохранения столичного городского совета, отделы здравоохранения соответствующих районных советов, Районный директорский совет южного Пешта, Республиканская служба скорой помощи и руководство заводских врачей.

- 2/ Одновременно пользуемся несколькими источниками извещения, и так об одном и том же подозрительном на инфаркт случае могут прибыть несколько извещений.

Мы пользуемся следующими источниками информации: учеты приемных отделений Республиканской службы скорой помощи, больниц и клинических учреждений, а также результаты вскрытия Юридического медицинского института и остальных патологанатомических институтов и отделов. Кроме этого пользуемся извещениями участковых врачей и данными журнала для записей умерших. Информацию о больных, которые на учете, Служба учета инфарктов приобрела в первую очередь путем активного розыска. На основе каждого извещения администраторы СУИ заполняют бланк "Извещение" /Приложение 1/, который, кроме паспортных данных больного, содержит данные об источнике извещения, о состоянии больного, о месте и времени госпитализации, в случае смертельного исхода — о времени наступления смерти.

О каждом извещенном врачи службы заполняют однотипный бланк "Основные сведения" /Приложение 2/, не позднее чем 28 дней после начала приступа. Кроме личных данных больного, бланк "Основные сведения" содержит данные о времени и месте приступа и медобслуживания, диагноз, а также критерии эпидемиологического диагноза.

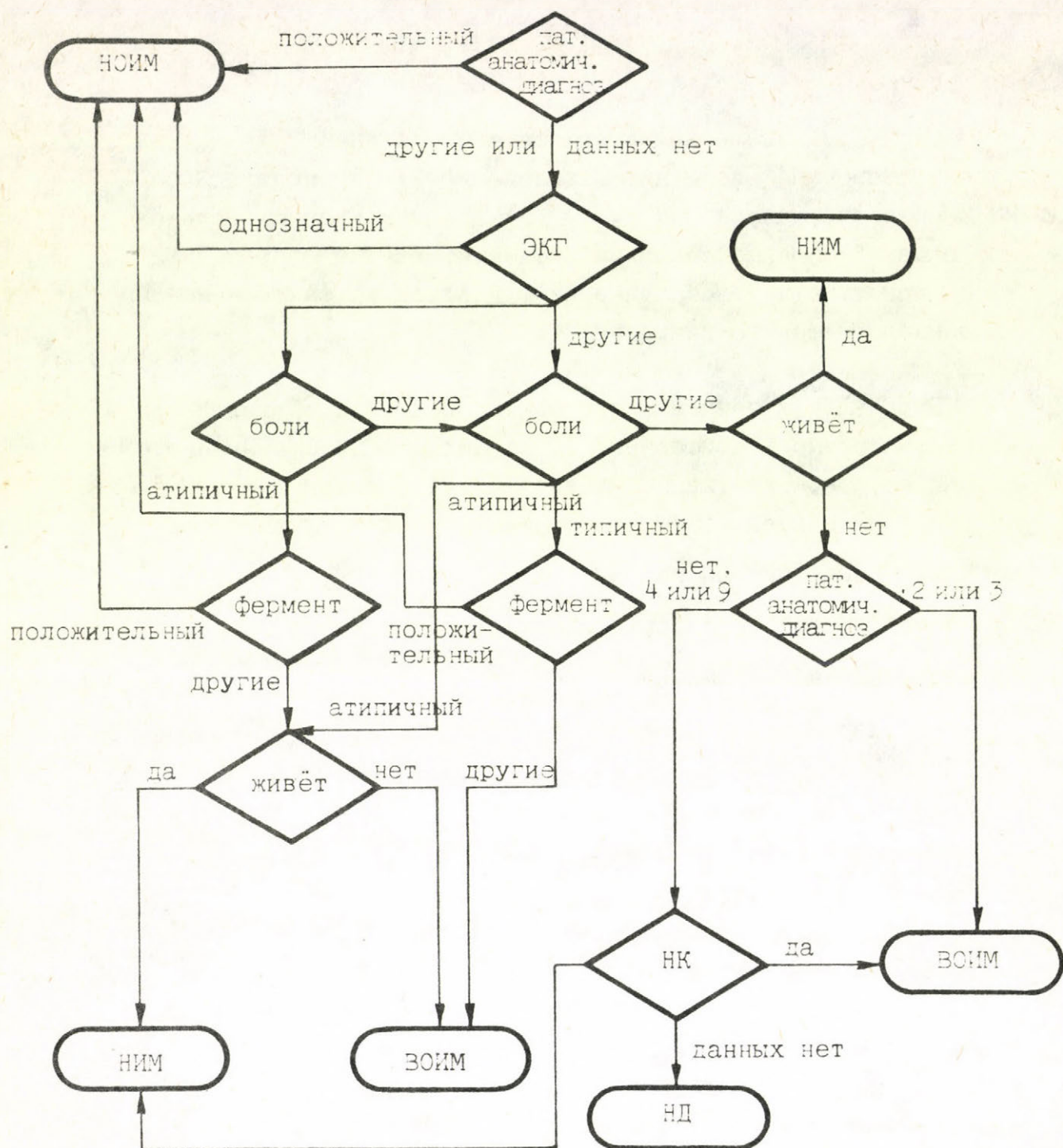
Определение эпидемиологического диагноза производится на основе четырех критериев: 1. характер боли в анамнезе, 2. ЭКГ, 3. ферментативные изменения, 4. патологоанатомический диагноз. На основе этих категорий диагноза следующие:

1. Несомненный острый инфаркт миокарда /НОИМ/;
2. Возможный острый инфаркт миокарда /ВОИМ/;
3. Неострый инфаркт миокарда /НИМ/;
4. Недостаточные данные /НД/
/Приложение 3/.

Путь определения диагноза демонстрируется диаграммой, указанной на стр. 7.

Для обеспечения полноты учета, каждые полгода администраторы службы учета пересматривают данные учетов больничных отделений и "журналов для записей скончавшихся" районных советов и сравнивают их с поступающими в службу учета данными. В результате этого пересмотра выявляется приблизительно 10 % новых случаев.

Число поступивших на учет сердечных приступов в течение одного года меняется в границах 2000-2500.



- НК - недостаточность кровоснабжения
- НОИМ - несомненный ОИМ
- ВОИМ - возможный ОИМ
- НИМ - неострый ИМ
- НД - недостаточные данные

ДИАГРАММА ОПРЕДЕЛЕНИЯ ЭПИДЕМИОЛОГИЧЕСКОГО ДИАГНОЗА

Непосредственно после создания, Служба учета инфарктов относилась организационно к Республиканскому кардиологическому институту. Опыт первых лет доказал, что ее деятельность подключается к работе территории, имеет региональный характер и поэтому с 1973 года служба учета является частью поликлиники столичной больницы имени Иштвана и относится к 1 терапевтическому отделению, где врачи службы выполняют и работу отделения.

Республиканский кардиологический институт помогает Службе учета инфарктов в обработке данных вычислительными машинами.

4. НЕОБХОДИМОСТЬ МЕХАНИЗАЦИИ УЧЕТА

Начатые в 1970 году измерения закончились в конце декабря 1971 года, а переосмотр - в 1972 году. Только после этого в 1973 году могли начать обработку данных и так результаты были в нашем распоряжении 4-5 лет после начала сбора данных. Так как учет инфарктов является оперативным учетом и имеет значение при принятии решений, медленность обратного получения результатов мешает более эффективной деятельности сети здравоохранения.

Сбор данных учета инфарктов немеханизированно - трудно- вато. Регулярное предоставление данных специалистам здравоохранения на территории, может быть решено только созданием громоздкого административного аппарата. Все это сделало необходимым автоматизацию учета инфарктов. Цель этого - путем регулярного предоставления данных - помимо информация руководителей здраво-

охранения, помочь участковым врачам в обслуживании инфарктных больных.

Служба учета инфарктов предоставляет информацию разным ступеням медицинского обслуживания и управления здравоохранения:

- 1/ участковым врачам - об их больных;
- 2/ больницам на территории Южного Пешта - главному врачу больницы и врачам - заведующим отделениями;
- 3/ главврачам Южно-пештских районов;
- 4/ предоставляет специальную информацию для научно-исследовательских работ.

5. ВЫВОДЫ

После десятилетней работы автоматизированной системы службы учета инфарктов, мы отметили следующие ее преимущества:

- 1/ В большой мере уменьшает администрацию учета;
- 2/ Обеспечивает постоянный контроль, улучшает точность учета;
- 3/ Делаеь более наглядным и доступным учет, чем помогает лучшему медицинскому обслуживанию;
- 4/ Благодаря регулярной обратной информации не только улучшает учет, а дает возможность более быстрому вмешательству;
- 5/ В случае надобности быстро может изготовить нужные информации вне установленного срока.

Приложение № 1

Учет острых инфарктов миокарда

Извещение

1. Место нахождения случая
1 = скорая помощь 2 = патологическая анатомия
3 = трибунал 4 = совет
5 = больница 6 = участковый врач
7 = другое
2. Состояние больного: 1 = живет 2 = умер
3. Ф. и. о. больного:
4. Пол: 1 = мужчина 2 = женщина
5. Адрес:
6. Возраст:
7. Дата /транспорт, смерть, госпитализация/:
год месяц день
8. Больница /где находится больной/:
9. Отделение: 1 = I-ое Терапевтическое 2 = II-ое Терапевтич.
3 = III-е Терапевтическое 4 = IV-ое Терапевтич.
5 = V-ое Терапевтическое 6 = VI-ое Терапевтич.
7 = Другое
10. Дата заполнения извещения:
год месяц день

Приложение № 2

Учет острых инфарктов миокарда

Основные сведения

Порядковый номер:

1. Центр: 01 = Дельпешт

2. Дата заполнения листа осмотра:

..... год месяц день

3. Ф. и. о. больного:

4. Пол: 1 = мужчина 2 = женщина

5. Ф. и. о. матери больного

6. Дата рождения: год месяц день

Порядковый номер:

7. Адрес:

8. Номер участкового врача:

9. Профессия:

Занятие:

Должность:

Место работы:

10. Приступ: 1 = был 2 = не был 3 = данных нет

11. Время начала приступа:

..... год месяц день час мин.

0 = не заполнено

9 = время начала приступа неизвестно

12. Состояние: 1 = живет 2 = умер

13. Наступление смерти: 1 = перед направлением в б-цу
2 = перед прибытием в б-цу
3 = после прибытия в б-цу
0 = живет

14. Время наступления смерти:

..... год месяц день час минута

0 = живет

9 = время наступления смерти неизвестно

15. Время прибытия в больницу:

..... год месяц день час минута

0 = не был в больнице

9 = неизвестный

16. Больница /в которую больной поступил/:

.....

17. Отделение /в которое больной поступил/

1 = I-ое Терапевтическое 2 = II-ое Терапевтическое

3 = III-ое Терапевтическое 4 = IV-ое Терапевтическое

5 = V-ое Терапевтическое 6 = VI-ое Терапевтическое

8 = другое

18. Проводилось ли лечение в интенсивном отделении?

1 = нет 2 = да 3 = неизвестно

Порядковый номер:

19. Поступление в интенсивное отделение:

..... год месяц день

0 = в интенсивном отделении не был

9 = неизвестно

20. Время выписки из больницы:

..... год месяц день

0 = в больнице не был

9 = неизвестно

21. Больница, из которой больной выписан:

.....

22. Отделение, которое выписало больного:

- 1 = I Терапевтическое 2 = II Терапевтическое
3 = III Терапевтическое 4 = IV Терапевтическое
5 = V Терапевтическое 6 = VI Терапевтическое
8 = Другое

23. Есть ли в анамнезе заболевание сердечной недостаточности?

- 1 = да, доказано 2 = да, недоказано
3 = нет 4 = неизвестно

24. Число инфарктов

- 0 = не было 1 = один
2 = два 3 = три
4 = четыре 9 = неизвестно

25. Время последнего инфаркта

- Год месяц день
0 = инфаркта не было
9 = время последнего инфаркта неизвестно

26. Оценка ЭКГ: 1 = однозначная 2 = неоднозначная 3 = прочая
4 = нормальная 5 = непредвиденная

27. Боли в грудной клетке: 1 = типичные 2 = нетипичные
3 = не были 4 = неизвестно

28. Сводная оценка ферментов: 1 = патологическая
2 = неоднозначна 3 = нормальная 5 = данных нет

29. SGOT 1 = патологическое 2 = неоднозначное
3 = нормальное 4 = данных нет

30. LDH 1 = патологическое 2 = неоднозначное
3 = нормальное 4 = данных нет

31. ISO 1 = патологическое 3 = неоднозначное
3 = нормальное 4 = данных нет

32. СРК 1 = патологическое 2 = неоднозначное
3 = нормальное 4 = данных нет

33. Клинический диагноз:

По мнению врача, констатирующего наступление смерти:
/если живет, на вопросы 34, 35, 36, 37 надо писать 0/

34. Причина смерти:

35. Основное заболевание:

36. Произведено ли вскрытие? 1 = да, 2 = нет

37. По данным вскрытия:

1 = свежая облитерация или атрофия /однозначно/

2 = старый рубец, облитерация или сужение больше
чем на 50%-ов /возможно/

3 = прочие знаки заболевания сердца в анализе
/вероятно/

4 = вышеприведенные данные отсутствуют

9 = неизвестно

38. Эпидемиологический диагноз:

1 = несомненный острый инфаркт миокарда

2 = возможный острый инфаркт миокарда

3 = неострый инфаркт миокарда

4 = недостаточные данные

ÖSSZEFOGLALÁS

In Hungary the number of the deceased in acute myocardial infarction /AMI/ amount to more than 10% of all mortality. This proportion as well as the frequency and incaliding effect of AMI in the active earning age accounted for the introduction of a system registering patients has begun in 1970 on the initiative of the WHO. After the completion of the examination the National Institute of Cardiology has continued this examination in shortened form. As a sequel to all these in 1977 a data base management system for a computer registry was organised. The aim of the system is to support by systematic data supply the information of leaders of public health, the district doctors in the care of infarction patients further to supply data for scientific research.

A cikk az 1977 óta működő számítógépes infarktus regiszter működésével, annak főbb részeivel foglalkozik.

НЕСКОЛЬКО ЗАМЕЧАНИЙ О РАБОТЕ ОТДЕЛА СТАТИСТИКИ
ИССЛЕДОВАТЕЛЬСКОГО ИНСТИТУТА ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И АВТОМАТИЗАЦИИ ВАН, СВЯЗАННЫХ С БИМЕДИЦИНСКИМИ ВОПРОСАМИ

д-р Иштван Ратко, д-р Михай Руда

В 1980 году в Ядвисине мы отчитались о деятельности Института - прежде всего нашего отдела -, которая связана с биомедицинскими вопросами.

В докладе мы хотим рассказать о продолжении упомянутых в Ядвисине вопросах.

На основании нашей научно-исследовательской работы, которую мы проделывали в связи с больничным морбидитизмом в Венгрии, сформировалась такая софтвер-техника, которая поддерживает программистов в обращении и в обработке сложных и больших статистических систем.

Упомянутая в Ядвисине Система Генера составляет основу этой софтвер-техники.

Наша теперешняя цель является внедрением в систему математических статистических задач.

Хотя статистические - даже сугубо биомедицинские носящие - программные пакеты существуют /например, БМДП/, но наша цель заключается в создании этих закрытых систем в более гибкой форме. Мы уже сделали доклады об этих планах и результатах /на международных конференциях в Венгрии/. По нашему мнению Система Генера, которая разработана в нашем отделе, является системой для генерирования программ, обеспечивает соответствующую базу для организации обработки больших и сложных систем данных.

Мы хотим использовать Систему Генера, кроме санитарных применений, и в исследовании растениеводческих систем.

Мы стремимся к дальнейшему развитию относительно мелких, удовлетворительных /местных требований/ с учетом болезни.

В настоящее время мы работаем над системной обработкой данных больных ожидающих операции на сердце.

Цели этой системы следующие:

1. Регистр больных ожидающих операции на сердце, списки о таких больных, которых удовлетворяют конкретные сложные условия. Эти списки помогают вызову на операцию больных.
2. Архивность соответствующих больных.
3. Предоставляет специальные информации, таблицы и статистики для научно-исследовательских работ и для врачей.

Система работает с терминалом, частично интерактивно.

Надеемся, что на следующем заседании мы подробнее изложим эту систему. Эта система осуществлена на ЭВМ Honeywell 66/20.

В этой области у нас есть новое представление о возможности использования персональной вычислительной машины, которые быстро распространяются в наши дни.

Цена этих машин все более снижается, но в то же самое время производительность и надежность машин растет.

В нашем отделе исследование машин типа TRS-80 опытное.

Биомедицинское применение персональной вычислительной машины очевидно. Например, потому что исследование биологических моделей по причине сложности требует интерактивной работы. А эта работа в большом вычислительном центре вообще довольно сложна.

Применение персональных вычислительных машин возможно и в больничном отделе, например, в случае приема больных и т.д.

В рамках сотрудничества СЭВ-а обрабатываются данные высокого давления крови среди молодежи. В другом докладе излагаются результаты фактора риска болезни /Янош Шолтес/.

ÖSSZEFOGLALÁS

A cikk az MTA SZTAKI számítógépes statisztikai osztályán folyó orvosbiológiai kutatások rövid áttekintését adja.

The paper gives a short survey of the Biomedicals researches of the Computer and Automation Institute Hungarian Academy of Sciences.

ВОПРОСЫ СИСТЕМЫ ИНФОРМАЦИИ ЛЕЧЕБНО-ПРОФИЛАКТИЧЕСКОЙ РАБОТЫ МЕДИЦИНСКОЙ СЛУЖБЫ

Автор д-р ШИМОН Пал, полковник медицинской службы

В период научно-технической революции вопросы увеличения числа информации и усложнения ее содержания становятся нерешенной проблемой. Эта же проблема существует и в медицинской службе. В прошлом качество и количество информации лечебной деятельности было иным, менее сложным, одним словом, врач сам способен был охватить, оценивать и сделать выводы с точки зрения установления диагноза. В наши дни задачи медицинской службы расширились не только в отношении лечебно-профилактической деятельности, но и в отношении других задач: хозяйственное обеспечение различных лечебно-профилактических учреждений, новые задачи противозидемического и гигиенического обеспечения, новые проблемы медснабжения /например, большое число медицинской аппаратуры и ее быстрая амортизация/, большие задачи по подготовке и усовершенствованию кадров и т.д. Таким образом, увеличение задач привело к еще большему возрастанию количества информации.

В области решения проблемы информации финансового-экономического характера, проблем обучения и усовершенствования были созданы системы и создаются в наши дни, а проблемами информации лечебно-профилактической деятельности вынуждена заниматься сама медслужба. Вопросы в этой области может поставить только сам врач: какая помощь требуется для правильной оценки информации лечения больных; и только после этого может оказать ему помощь организатор системы, программист, оператор и т.д. Мы можем быть уверены в одном: правильная оценка информации лечебно-профилактической работы возможна только с применением ЭВМ и в этой очень важной работе можно достичь успеха только при тесном сотрудничестве врача и организатора. Общеизвестно, что в медслуж-

бах всего мира ведется работа по использованию ЭВМ для оценки информации, автоматизации и управления медслужбой, но главным образом, в рамках лечебных учреждений /больницы, госпитали, поликлиники/. Мы хотели бы решить этот вопрос в области оказания первичной медицинской помощи: а именно там, где больной в большинстве случаев впервые встречается с врачом - в кабинете участкового врача. Естественно, мы занимались только лечебно-профилактической стороной вопроса и не затрагивали других проблем. /Схема № 1, см. на следующей странице./

ПЕРВЫЙ ЭКСПЕРИМЕНТ

Вначале мы создали такую систему, где исходным документом являлась амбулаторная книжка участковых врачей. Этот документ содержит случаи посещения, заболевания, анамнез, состояние больного, то есть статус, длительность заболевания, а также диагноз и разнообразные результаты процесса лечения. Амбулаторная книжка все эти важные элементы информации лечебно-профилактических работ содержит таким образом, что включает в себя результаты лечения в госпитале, санатории и других учреждениях медслужбы, то есть начало и конец лечебного процесса. Правда, эти данные не подробные, но отражают суть дела /например, конечный диагноз заболевания/.

Мы создали вторичный документ - носитель данных - для обработки на ЭВМ. Этот документ заполняется вручную, одна строка содержит данные одного заболевания. Контроль данных производится на месте и в вычислительном центре: ручным способом, логически и машинным способом. Подготовка данных для обработки на ЭВМ происходит с помощью перфорации; сама обработка автономная, то есть офф-лайн.

Результаты обработки - выпуск, то есть аутпут - состоят из таблиц, которые отличаются друг от друга по уровням управления медслужбы на разных этапах. Кроме того эти таблицы - с помощью специальных параметров заболеваемости и т.д. - содержат

Схема № 1.

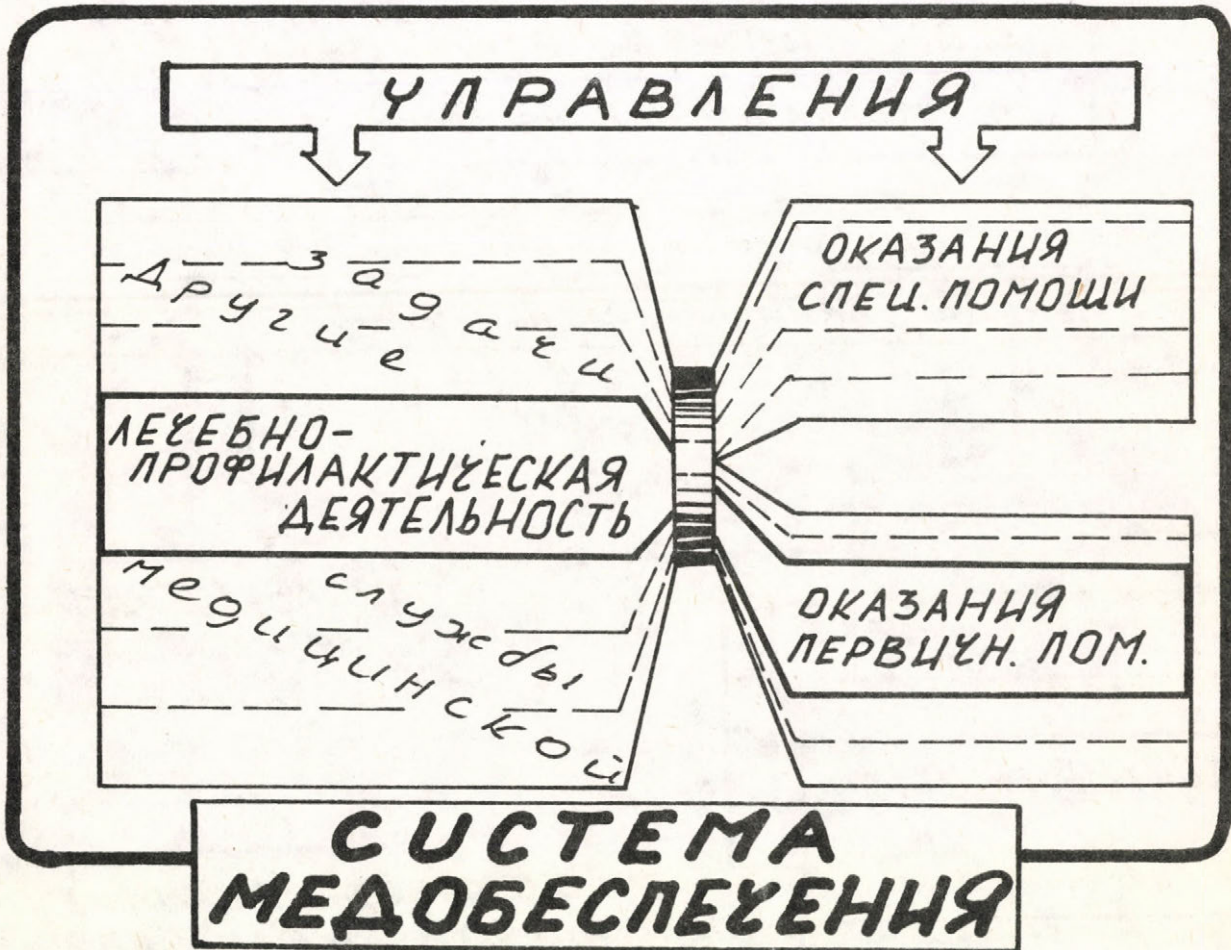
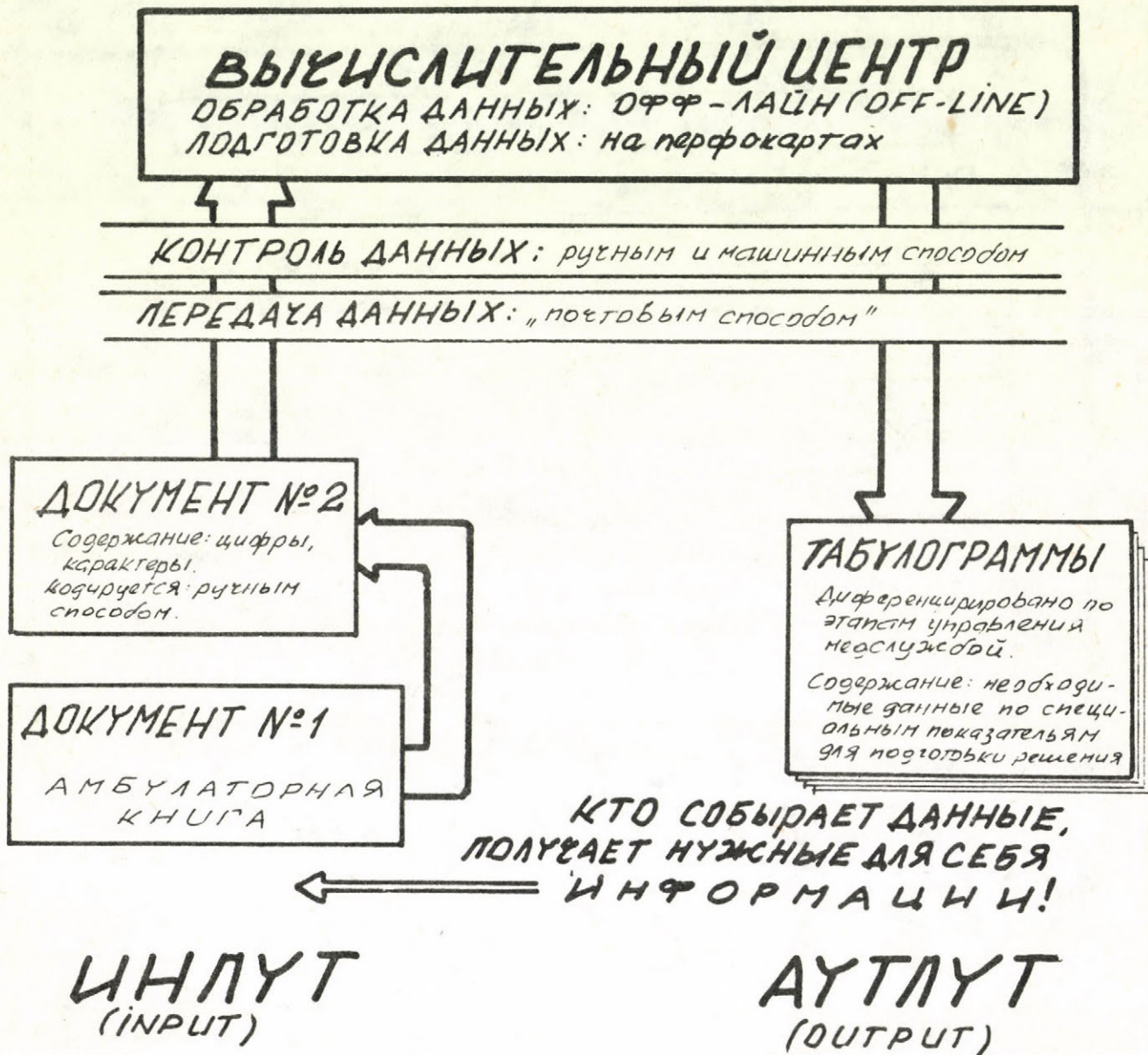


Схема №.2.



только самые важные предварительные информации для принятия решения- Это значит, что управление медслужбой получает данные только по интересующим ее вопросам. Кроме того, участковый врач от которого поступают данные, сам получает важную для себя информацию и освобождается от ручных статистических донесений. Таким образом он /участковый врач/ также заинтересован в правильной обработке информации. Естественно, это также относится и к разнообразным уровням управления медслужбой. Эта система представлена на схеме № 2. /См. схему на следующей странице./

Из этой работы можно сделать уже несколько практических выводов:

- медицинские кадры привыкли к этому методу работы и они уже требуют таблограммы и службу ЭВМ, руководство на различных уровнях имеет возможность получения более ценной информации;
- выяснилось, что нужно создать банк данных для медицинских информаций и необходимо найти возможность для каждого больного в отдельности сохранить самые важные информации об изменении состояния его здоровья;
- в дальнейшем надо иметь ввиду требования научно-медицинской работы, надо иметь ввиду связь между различными этапами лечебно-профилактической деятельности и вертикально /участковый врач - врач-специалист госпиталя/ и горизонтально /отрасли лечебной работы - хирургия, терапия и т.д. и различные стороны профилактической работы - осмотр и диспансеризация/.

Принимая во внимание указанный опыт, мы приступили к созданию новой системы.

ВТОРОЙ ЭКСПЕРИМЕНТ

Мы ясно представляли стоящие перед нами задачи:

- самое важное создать систему постоянного наблюдения за сос-

тоянием здоровья человека /больного/, то есть создать систему ориентированную на больного и одновременно сохранить и расширить результаты системы управления медслужбой;

- необходимо было отобрать самые важные данные процесса лечения, процесса профилактической работы, изменения состояния здоровья человека и определить необходимые для этого самые важные документы;
- хотя эта система разрабатывается для первичного медицинского этапа, надо было иметь ввиду и самые важные узлы связи между другими этапами медицинской помощи /госпитали, больницы/;
- естественно, что при составлении банка данных надо было иметь ввиду различные способы доступности данных /он-лайн, офф-лайн/, различные этапы хранения данных, например, с точки зрения научной работы.

Проблемы создания ин-пут: имея ввиду основные стороны лечебно-профилактической работы, так называемая лечебная работа, так называемая диспансеризация и контроль состояния здоровья и предполагая замкнутость системы, вначале мы создали группы информации - персональные данные, данные о работе и образе жизни, анамнестические данные /личный, семейный и т.д./, частота обращения к медслужбе, данные осмотра, диагнозы и осложнения, решения врача /например, на каком этапе, каким способом надо лечить больного/, этапы медицинской службы /где протекает лечебно-профилактическая деятельность/, данные исхода, результаты лечебно-профилактической деятельности.

Как показывает или стремиться показать группирование данных, верная информация получается только тогда, если мы считаем лечебно-профилактическую деятельность постоянным процессом, на всем протяжении которого следует стремиться собирать данные. Другими словами, мы поставим "монитор" для наблюдения за состоянием здоровья человека и получим своевременно данные о его

заболевании, о работе лечебного учреждения и можем получить правильные исходные данные для научной работы.

Документы: основная карточка, документы для донесения о заболевании, диспансеризации и контроле.

Подготовка данных для машинной обработки: например, с помощью оптического устройства считывания /в этом случае контроль централизован/ и использование перфокарт и перфолент ВИДЕОПЛЕКС /в этом случае контроль на месте/.

Способ обработки данных определяется по типу службы аутпут.

Проблемы создания аутпут: здесь не учитываем проблемы научной работы и остановимся только на двух сторонах обработки данных, то есть на двух сторонах аутпут: система ориентирования на больного и система управления медслужбы.

Система ориентирования на больного обеспечивает так называемый "медицинский паспорт", немедленная достигаемость он-лайн, который содержит самые важные данные о человеке. Другие документы - достигаемость офф-лайн - следующие: донесение о диспансеризации лиц и донесения о контроле, а также донесения о заболевании /эпикриз/. Кроме этого планируется так называемая история состояния здоровья человека, которая частично может содержать различные заболевания, отклонения от нормы. Следует подчеркнуть, что эти аутпуты являются настоящими продуктами машинной обработки.

Система управления медслужбой: ее основная структура не изменилась, но и содержание стало более плотным и повысилась ценность подготовки принятия решения.

Мы отлично знаем, что кроме вышесказанных при создании системы нам надо иметь ввиду следующее: основное внимание обратить на связи между различными этапами медслужбы, нельзя считать зада-

чу законченной, но даже более того - нам необходимо иметь ввиду постоянное дальнейшее совершенствование. Мы думаем, что эта система позволит нам создать фундамент современной информации лечебно-профилактической деятельности, к которой можно присоединить уже существующие подсистемы /например, программа анализа ЭКГ/ или подсистемы, которые будут подготовлены в будущем.

Планирование второго эксперимента окончено, начали работу программисты. Основная машина ЭСР-40 /можно ЭСР-55/. Система будет готова вероятно к концу следующего года. Большое спасибо за внимание.

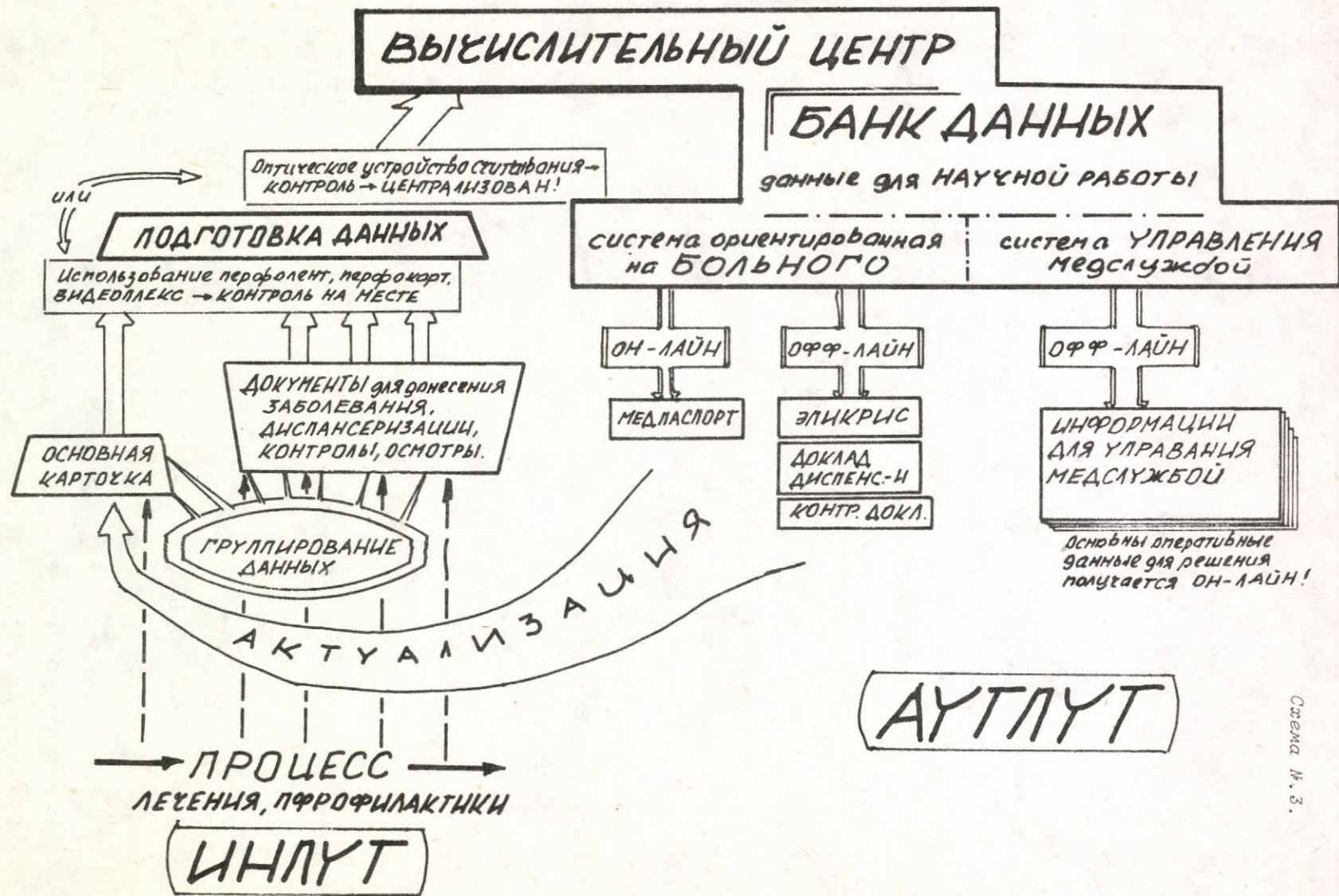


Схема № 3.

ÖSSZEFOGLALÁS

AZ EGÉSZSEGÜGYI SZOLGÁLAT GYÓGYÍTÓ-MEGELŐZŐ TEVÉKENYSÉGE INFORMÁCIÓS RENDSZERÉNEK KORSZERŰSÍTÉSI PROBLÉMÁI

Simon Pál

Az egészségügyi szolgálat alaprendeltetését alkotó gyógyító-megelőző ellátás hatékonyságának növeléséhez korszerű számítástechnikai módszereket, eszközöket is alkalmazó információs rendszer kell. A bemutatott és működő vezetésorientált számítógépes információs rendszer az alapellátásra támaszkodik és minden szintű szakmai vezetőt támogat a szakmai irányításban a legszükségesebb morbiditási, betegforgalmi adatokat, viszonyszámokat tartalmazó döntéselőkészítő információkkal. Lényeges mozzanat, hogy az adatszolgáltató is visszakap feldolgozott információkat.

A fejlesztés lényege: a vezetés orientáltság mellegg a betegorientáltságú - az egyén egészségi állapotát aktuálisan nyomonkövető - információs rendszer kialakítása az alapellátásra támaszkodóan a szakellátás felé tartó fejlesztési iránnyal.

SUMMARY

ON HOW TO MAKE MORE UP TO DATE THE INFORMATION SYSTEM OF MEDICAL AND PREVENTIVE CARE IN HEALTH SERVICE

It is necessary to have proper information system including the usage of up to date computer methods and equipments to improve the efficiency of medical and preventive care which create the basis of health service.

The demonstrated and working management-oriented computerized information system bases on District Health Service and helps professional manager on every level by producing information on morbidity and patient circulation, which help to make decision.

It is important step that the date-supplier can get back processed information.

Summary of our results: the development of a patient-oriented system besides the management-oriented one - which can follows the actual health condition of an individual - basing on District Health Sercie.

It is obviously means an advance and help for the District Special Service as well.

SOME COMPUTER-AIDED SYSTEM FOR THE MINISTRY OF HEALTH

Benedek Srajber

Central Statistical Office
Computer Service for State Administration

Main profile of the *Application Department*:

developping information systems for some ministries to give help to administration and to dicision making.

In connection with the *Ministry of Health* we have three important fields:

- I. Management systems of health.
- II. Patient care systems.
- III. Methods and methodologies to evaluation of medical data.

I. Management systems of healt

1. Complex system for the investigation of morbidity in Hungary (KOMOV).

Data base (90%, HwB 66/20, IDS I, national survey).

Beginning: 1978

Working system: 1981

2. National Information System of Child Health (OGYIR)

Beginning: 1979

Model-experiment in district guided by the Ministry of Health (Dept. IV.)

Working model: 1982

3. Model-experiment: a computer aided mass screening system of pulmonary diseases

Beginning: 1979

Working system: 1982

II. Patient care systems

1. Computer-aided Planning System in Radiotherapy.
Since 1979: Nationwide Working system with six terminals.

2. Hospital Information system (model in military hospital)

Beginning: 1978

Working: 1980

3. National Registration and Analysing System in Children's Leukaemia

Working System 1979

(Interactiv programs analyse the survival tables with the logrank test.)

III. Methods and methodologies to evaluation of medical data

Methods:

1. Program system for the multicatorial investigation of national surveys (PERINAT)

Developed: 1972-80

First application of large size: The investigation of the new-born population in Hungary (representative sample: 30.000 cases).

Figure 1.

2. Programs for analysis of psychological tests.
 - a.) Rorschach test
 - b.) Szondi test
3. WHO standards on computer HwB
4. Creating unified code-system in psychology and interval medicine in cooperation with the corresponding institute of health.

12.05.1981.

ÖSSZEFOGLALÁS

Benedek Srajber

Az előadás olyan egészségügyi alkalmazásokról ad tömör tájékoztatást, amelyek az Egészségügyi Minisztérium vagy valamely országos egészségügyi intézmény megbízása, ill. közreműködése alapján realizálódtak.

СОДЕРЖАНИЕ

Доклад кратко информирует о таких применениях, которые были реализованы на основе заказа и с участием Министерства Здравоохранения или некоторого всевенгерского учреждения здравоохранения.

AN EFFICIENT SOLUTION OF BANDED TOEPLITZ
SYSTEMS BY APPLICATION OF FAST FOURIER TRANSFORM *

G.Beck

Research Institute for Applied Computer Sciences

ABSTRACT

Banded Toeplitz matrices of large size occur in many practical problems [1]-[4]. Here the problem of solving linear equations of the type $T\underline{x}=\underline{y}$, when T is a banded Toeplitz matrix, is considered. It is shown by circular extension of T , that such equations may be solved in $O(n\log_2 n)$ rather than $O(n^2)$ computations as in recently applied algorithms. The advantage of the presented algorithm is realised in those applications where the coefficient matrix of the linear equations is a large scale, sparse matrix.

I. INTRODUCTION

In this paper we consider the problem of solving linear systems, i.e. such sets of linear equations, where the coefficient matrix is a quadratic, banded Toeplitz one. Such matrices occur in many applications. Some essential physical and mathematical application fields are [1]-[4]:

- image processing
- digital filtering
- signal processing
- solution of certain partial differential equations.

In a large number of these situations it is desirable to solve for \underline{x} , a set of linear equations

$$T\underline{x}=\underline{y}$$

(1)

* Ez az előadás nem a KNVVT-n hangzott el.

where T is an $n \times n$, banded Toeplitz matrix.

The well-known Gauss elimination algorithm [5] and the recursive algorithm of Lenison, Trench and Zohat [7]-[9] (henceforth abbreviated L-T-Z algorithm) have been recently used for solving problems of type (1). The number of computations required is $O(n^3)$ and $O(n^2)$, respectively for solving system (1) by Gauss and L-T-Z algorithms. (One computation is defined as one multiplicative and one additive operation).

In the following we propose an improved method, based on the circular extension of coefficient matrix, and on application of fast Fourier transform algorithm where the number of computations required are reduced to $O(n \log_2 n)$. The advantage of the method is realized in large scale, sparse Toeplitz systems.

K. Jain published an algorithm for the solution of banded Toeplitz systems, based on circular extension in 1978 [9], but this algorithm is suitable only to solve even more special systems than (1). The method of this paper worked out by extension of Jain's algorithm.

Since the fast Fourier transform algorithm [10] (henceforth abbreviated FFT) place an essential role in increasing efficiency, we may regard our method as an interesting practical application of FFT algorithm.

II. SOME BASIC DEFINITIONS AND NOTATIONS

We start with a brief note regarding notation. We use upper-case letters for matrices, lower-case letters for scalars and underscored lower-case letters for vectors. We use the following notes and abbreviations:

- we denote component-by-component multiplication of two vectors with: *
- we denote component-by-component division of two vectors with: *
- *

- the discrete Fourier transform of the vector \underline{a} is:
DFT(\underline{a})
- the inverse discrete Fourier transform of \underline{a} is: IDFT(\underline{a})
- the log denotes the logarithm based on 2, that is $\log n$
means $\log_2 n$.

Definition:

A matrix T whose ijth element t_{ij} is a function of (i-j) (rather than of i,j separately), is called a Toeplitz matrix i.e. for an nxn Toeplitz matrix

$$t_{ij} = t_{i-j} \quad 0 < i, j \leq n$$

Definition:

An nxn matrix C is called circulant or cyclic matrix, when

$$c_{ij} = c_{i-j} \pmod{n} \quad 0 < i, j \leq n$$

that is

$$c_{ij} = c_{i-j \pm kn} \quad (k=0,1,2,\dots)$$

III. THEORETICAL DEVELOPMENT

Consider the set of linear equations

$$T\underline{x} = \underline{y} \tag{1}$$

where T is an nxn, nonsingular, banded Toeplitz matrix, i.e.

$$\begin{aligned} t_{ij} &= t_{i-j} & 0 < i, j \leq n \\ \exists p, q < n & \quad t_q \neq 0, \quad t_{-p} \neq 0 & \text{ and} \\ t_k &= 0 & \text{ for every } k > q \text{ or } k < -p \end{aligned}$$

The matrix T has the form:

$$T = \begin{bmatrix} t_0 & t_{-1} & \dots & t_{-p} & 0 & \dots & 0 \\ t_1 & & & & & & 0 \\ \vdots & & & & & & \vdots \\ t_q & & & & & & 0 \\ 0 & & & & & & t_{-p} \\ \vdots & & & & & & \vdots \\ 0 & \dots & 0 & t_q & \dots & t_1 & t_0 \end{bmatrix}$$

In the following we will show, that the solution of the Topelitz equation (1) can be reduced to the solutions of two sets of linear equations. One set requires solving a set of linear equations, where the coefficient is an $n \times n$ circulant matrix, which can be accomplished in $O(n \log n)$ computations via the FFT. The other set requires solving a much lower rank $O(p+q)$ of Toeplitz equations which can be accomplished in $O((p+q)^2)$ computations by the L-T-Z algorithm.

Let T_c denote the circular extension of matrix T , i.e. an $n \times n$ circulant matrix which contains the elements of T . Matrix T_c will be the following:

$$T_c = \begin{bmatrix} t_0 & t_{-1} & \dots & t_{-p} & 0 & t_q & \dots & t_1 \\ & t_1 & & & & & & \vdots \\ & & & & & & & t_q \\ & & & & & 0 & & \\ & & & & & & & 0 \\ & & & & & & & t_{-p} \\ & & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & t_{-1} \\ & & & & & & & t_0 \\ t_{-1} & \dots & & t_{-p} & 0 & t_q & \dots & t_1 \end{bmatrix} \quad (2)$$

Note: The circular extension of matrix T exists if $p+q+1 \leq n$.
 When this condition is false, i.e. T isn't sufficiently sparse, the L-T-Z algorithm is used.

T_b is defined by the equation:

$$T_b = T_c - T \quad (3)$$

Using (1), (2) and (3) and rearranging the terms we can write:

$$T_c \underline{x} = \underline{y} + T_b \underline{x} \quad (4)$$

Now, let us assume that T_c is also nonsingular. So equation (4) can now be rewritten as

$$\underline{x} = T_c^{-1} \underline{y} + T_c^{-1} T_b \underline{x} \quad (5)$$

And with the notations:

$$B = T_c^{-1}, \quad \underline{z} = T_c^{-1} \underline{y} \quad (6)$$

(5) becomes

$$\underline{x} = \underline{z} + B T_b \underline{x} \quad (7)$$

Let us consider the product $T_b \cdot \underline{x}$ on the right side of equation (7):

$$T_b \cdot \underline{x} = \begin{bmatrix} & & & & t_q & \dots & t_1 \\ & & & & & & \vdots \\ & & & & & & t_q \\ & & & 0 & & & \\ & t_{-p} & & & & & \\ & \vdots & & & & & \\ t_{-1} & \dots & t_{-p} & & & & \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \end{bmatrix}$$

Let F and G denote the $q \times q$ upper and the $p \times p$ lower triangle matrices in T_b , furthermore $\underline{x}^i, \underline{x}^t$ the $p \times 1$ and $q \times 1$ column vectors containing p initial $\{x_0, x_1, \dots, x_{p-1}\}$ and q terminal values $\{x_{n-q}, \dots, x_{n-1}\}$ of \underline{x} , respectively.

In this notations the former product is:

$$T_b \underline{x} = \begin{bmatrix} & & F \\ & 0 & \\ G & & \end{bmatrix} \cdot \begin{bmatrix} \underline{x}^i \\ \underline{x}^m \\ \underline{x}^t \end{bmatrix} = \begin{bmatrix} F \cdot \underline{x}^t \\ 0 \\ G \cdot \underline{x}^i \end{bmatrix} \quad (8)$$

We further define \underline{u} and \underline{v} column vectors, as

$$\underline{u} = F \underline{x}^t, \quad \underline{v} = G \underline{x}^i \quad (9)$$

Equation (7) can now be rewritten as

$$\underline{x} = \underline{z} + B T_b \underline{x} = \underline{z} + B \cdot \begin{bmatrix} \underline{u} \\ 0 \\ \underline{v} \end{bmatrix} \quad (10)$$

If the values of \underline{u} and \underline{v} were known, (10) could be evaluated easily to give \underline{x} .

Since \underline{u} and \underline{v} depend on \underline{x} , (10) as such is an implicit equation. If we partition the matrix B, as

$$\begin{array}{cccc}
 & B^{ii} & B^{im} & B^{it} & p \\
 B = & B^{mi} & B^{mm} & B^{mt} & n-p-q \\
 & B^{ti} & B^{tm} & B^{tt} & q \\
 & q & & p &
 \end{array} \quad (11)$$

then it is easy to obtain, from (9) and (10), the identities

$$\begin{aligned}
 G^{-1} \underline{v} = \underline{x}^i &= B^{ii} \underline{u} + B^{it} \underline{v} + \underline{z}^i \\
 F^{-1} \underline{u} = \underline{x}^t &= B^{ti} \underline{u} + B^{tt} \underline{v} + \underline{z}^t
 \end{aligned} \quad (12)$$

where \underline{z}^i and \underline{z}^t are the $p \times 1$ and $q \times 1$ initial and terminal vectors of \underline{z} . The matrices F and G are nonsingular since t_{-p} and t_q in T are nonzero.

By rearranging (12) we can write

$$\begin{array}{ccc}
 & -B^{tt} \underline{v} + (F^{-1} - B^{ti}) \underline{u} = \underline{z}^t \\
 (G^{-1} & -B^{it}) \underline{v} + & -B^{ii} \underline{u} = \underline{z}^i
 \end{array} \quad (13)$$

Let

$$\begin{array}{ccc}
 0 & F^{-1} & q \\
 J_1 = & G^{-1} & 0 & p \\
 p & q & &
 \end{array} , \quad \begin{array}{ccc}
 B^{tt} & B^{ti} & q \\
 J_2 = & B^{it} & B^{ii} & p \\
 p & q & &
 \end{array} \quad (14)$$

and

$$J = J_1 - J_2 \tag{15}$$

Therefore equations (13), when written in a $(p+q) \times (p+q)$ matrix form, become

$$J \cdot \begin{matrix} \underline{v} \\ \underline{u} \end{matrix} = \begin{matrix} \underline{z}^t \\ \underline{z}^i \end{matrix} \tag{16}$$

Since equation (16) is a set of linear equations and since J is nonsingular (it has to be true hence we supposed that the matrix T is nonsingular and so (12) must be $(p+q)$ linearly independent equations in $(p+q)$ unknowns $(x_0, x_1, \dots, x_{p-1}, x_{n-q}, \dots, x_{n-1})$), (16) can be solved to obtain the values \underline{u} and \underline{v} .

Therefore the overall algorithm to obtain the solution of system (1) has 5 simple steps:

1. calculate B by (6)
2. calculate \underline{z} by (6)
3. solve the $(p+q)$ equations of (16) to obtain \underline{u} and \underline{v}
4. calculate the product $B \cdot \begin{matrix} \underline{u} \\ \underline{v} \end{matrix}$
5. evaluate the solution vector \underline{x} by (10)

If we used for each step of the former method the general algorithms, i.e. for the inversion of matrix T , and for solution of linear equation (16) the well-known Jordan and Gauss elimination algorithms, and for the matrix-vector multiplication the traditional method, then the number of computations required for the solution of (1) would be much greater than, for L-T-Z algorithm. However by building on the special properties of circulant matrices and the application of

FFT algorithm, the number of computations required can decrease significantly.

For the reduction of computations we need the following theorems

Theorem1(T1): The inverse of a nonsingular circulant matrix is also a circulant matrix. (Proved in [11])

Theorem2(T2): Suppose that C is an $n \times n$ circulant matrix and D is the inverse of C . Let \underline{c} and \underline{d} denote the first column of C and D , respectively. Then

i, $\underline{d} = \text{IDFT}(\underline{1} * \text{DFT}(\underline{c}))$

ii, by the application of the FFT, D can be obtained with

$$2n \log n + n$$

computations. [12]

Theorem3(T3): Let \underline{c} denote the first column of an $n \times n$ circulant matrix C , \underline{a} is a column vector with n components and \underline{b} is the product of C and \underline{a} . Then

i, the product vector \underline{b} can be obtained by the following formula:

$$\underline{b} = \text{IDFT}(\text{DFT}(\underline{c}) * \text{DFT}(\underline{a}))$$

ii, the calculation of \underline{b} can be accomplished in

$$3n \log n + n$$

computations via the FFT. [12]

Theorem4(T4): Matrix J , in formula (15) is a Toeplitz matrix and its elements can be calculated by

$$\frac{p^2 + q^2}{2} + 0 \quad (p+q)$$

computations. [12]

Based on these theorems every step of the former 5 step method can be reformulated:

1'. The calculation of matrix B:

Suppose that \underline{t} is the first column of T_c , then, according to T1, B is also a circulant matrix and its first column, by T2 is:

$$\underline{b} = \text{IDFT}(\underline{1} * \text{DFT}(\underline{t}))$$

computations required: $2n \log n + n$

Since B is a circulant matrix its first column defines it completely.

2'. The calculation of vector \underline{z} :

According to T3, the product-vector \underline{z} can be obtained by the formula:

$$\underline{z} = \text{IDFT}(\text{DFT}(\underline{b}) * \text{DFT}(\underline{y}))$$

computations required: $3n \log n + n$

3'. The solution of the $(p+q)$ equations of (16):

According to T4, the coefficient matrix of the linear equation (16) is a $(p+q) \times (p+q)$ Toeplitz matrix, and so its solution can be obtained by the L-T-Z algorithm via $3(p+q)^2$ computations.

4'. The calculation of the product $B \cdot \frac{\underline{u}}{\underline{v}}$:

According to T3, for this multiplication we need only $3n \log n + n$ computations.

5'. The evaluation of the solution vector \underline{x} :

In this step we must add vector \underline{z} and the product-vector obtained in the step 4.

On the basis of these it can be seen, that the number of computations required for the presented method is

$$O(n \log n) + O((p+q)^2)$$

and it will be efficient in comparison with other methods, when the coefficient matrix is sufficiently large scale and sparse, i.e. $p, q \ll n$. (In the applications referred in the introduction, such large scale, sparse matrices occur.)

Note: It seems to be a difficult problem to decide whether the circulant matrix T_c is singular or not. However we have a simple method for testing whether a circulant matrix is nonsingular.

Theorem5 (T5): A circulant matrix is nonsingular if and only if the discrete Fourier transform vector of its first column doesn't contain any zero component.
(Proved in [12])

IV. ABOUT THE IMPLEMENTATION

Since we implemented the presented algorithm in Fortran language [13], so we were able to test the mentioned algorithms on an IBM 370/145 type Computer.

In the following table the rounded CPU times required to solve various sets of linear equations with banded Toeplitz matrices can be seen.

N (p=q=12)	Gauss elimination		L-T-Z algorithm		Proposed algorithm	
	min	sec	min	sec	min	sec
32		4		1		1
128	2	59		9		2
256	23	11		32		3
512			2	01		6
1024			8	03		11
2048			32	35		21
4096			127	33		44
8192					1	25

ACKNOWLEDGEMENT

I should like to express my thanks to Pál Sebestyén for raising this subject and for the valuable support in the solution of the problem.

I also thankful to László Simonfai and Imre Szalay for their useful comments and suggestions.

REFERENCES

- 1 A.K. Jain: An operator factorisation method for restoration of blurred images
IEEE Trans. Comput. pp 1061-1071, Nov. 1977
- 2 D.C. Farden and L.L. Scharf: Statistical design of digital filter
IEEE Trans. Acoust., Speech, and Signal Processing pp. 188-196. June 1970.

- 3 D. Preis: The Toeplitz matrix: its occurrence in antenna problems and a rapid inversion algorithm
IEEE Trans. on Antennas and Propagation pp. 204-207,
March 1972
- 4 E. Angel and R. Bellman: Dynamic Programming and Partial
Differential Equations
New York: Academic, 1971.
- 5 F. Móricz: Numerical Analysis II.
Budapest, Educational Publishing House 1975 /in
Hungarian/
- 6 N. Levinson: The Wiener RMS /Root mean square/ error
criterion in filter design and prediction
J. Math. Phys., pp. 261-278, Jan. 1947
- 7 W. F. Trench: An algorithm for the inversion of finite
Toeplitz matrices
J. SIAM, pp. 515-522, Sept. 1964.
8. S. Zohar: The solution of a Toeplitz set of linear
equations
J. ACM, pp. 272-276, Apr. 1974.
- 9 A.K. Jain: Fast inversion of banded Toeplitz matrices by
circular decompositions
IEEE Trans. Acoust., Speech, and Signal Processing
pp. 121-126, Apr. 1978.
- 10 G.Beck: The fast Fourier transform, with an application
Diploma work
Szeged, Attila József University, 1978 /in Hungarian/
- 11 R.M.Gray: Toeplitz and circulant matrices: A review
Stanford Univ., Stanford, CA, Tech.Rep., 1970.
- 12 G.Beck: A fast algorithm for the solution of banded
Toeplitz sets of linear equations Applied Math.
J. Hung. Acad. Sci., 1982 /in Hungarian/

- 13 G.Beck: Fast algorithm for the solution of linear equations with banded Toeplitz matrices
SZÁMKI Research paper, Budapest, 1980. /in Hungarian/

ÖSSZEFOGLALÁS

Nagyméretű Toeplitz-típusú szalagmátrixok igen sok gyakorlati alkalmazási területen felmerülnek [1]-[4]. A cikkben olyan $T\underline{x}=\underline{y}$ alakú lineáris egyenletrendszerek megoldásával foglalkozunk, melyek együttható - mátrixa Toeplitz-típusú szalagmátrix. Megmutatjuk, hogy a T cirkuláris kiterjesztésével az ilyen egyenletrendszerek - a jelenleg alkalmazott $O(n^2)$ műveletigényű eljárásokkal szemben - $O(n \log_2 n)$ művelettel megoldhatók. Az eljárás abban az esetben nyújt hatékonyságnövekedést, ha a megoldandó lineáris egyenletrendszer együtthatómátrixa nagyméretű és ritka.

