18.1-04 SIMULTANEOUS OPTIMIZATION FOR AUTOMATIC DIFFRACTOMETER AND CRYSTAL STRUCTURE DETERMINATION.

By V.Debelle, A.Hardy , Laboratoire de Cristallochimie Minérale , J.Nérault , Laboratoire de Dynamique des Fluides, Université de Poitiers, 40 av. Recteur Pineau F-86022 Poitiers Cédex,France et M.Tournarie DPhG/SRM, CEN de Saclay, F-91190 Gif sur Yvette, France.

By the use of TOURNARIE's criterium (Cybernetica (1978) 31, 227-262) the suppression of human intervention is obtained on the one hand to carry out the physical measurement for automatic diffractometer and , on the other, to calculate the parameters of crystal structure by trial and error methods. The decisions are better informed of a broader scope . The simultaneousness of the two operations , data collection and processing , which have been up to now successive , allow a gain in time and in efficiency, since each piece of information is immediately exploited . For an equal degree of accuracy , the duration of the measurement as well as of the structure determination is divided by 3 .

The verification of the increase in accuracy and the generalization to any other couple, automatic apparatus - indirect non-linear system of unknowns , is being envisaged .

18.1-05 The NRC PDP-11 CRYSTAL STRUCTURE SYSTEM By E.J. Gabe and F.L. Lee, Chemistry Division, National Research Council of Canada, Ottawa, Ontario, K1A 0R6, Canada.

The system which is converted from the NRC PDP-8e crystal structure system (Larson and Gabe, Computing in Crystallography, 1978) will perform all the standard computation of crystal structure analysis. It should run on any PDP-11 machine with a minimum of 32K words task space under RSX-11M. It is written entirely in Fortran for the IV PLUS compiler. In addition about 10 MB of disc space are needed (5MB for sources etc and 5MB for task modules and working files).

The RSX-11M operating system imposes a 32K task size limit on its users, and the surprising fact that emerged from the conversion is that the saving in memory usage, 16-bit instead of 12 bit words, and 2 words instead of 3 words per variable, is more than offset by the size and complexity of the link to the time-shared operating system. Thus it was not possible to increase the size of the problem which could be handled by the apparently larger computer. With fast floating point hardware on both machines (FPII-EA and FPP-12) the system runs between 5 and 6 times faster on the 11/60 than on the 8e.

The system is based on two major files. The crystal data (.CD) file contains cell, symmetry, scattering factor and atomic parameter data. The reflection file (.RE) contains all data pertinent to reflections.

Routines in the package are:

CDFILE - To create a .CD file.
ABSORP - To calculate transmission factors and β values by Gaussian integration.
DATRD2 - To reduce intensity data from single or multiple sets to Fo and E values, and create the reflection file (.RE).
CDEDIT - To modify the CD file by changing parameters, inserting atoms etc.
FOURR - To calculate all types of Fourier maps.
UNIMOL - To calculate distances from positions in a CD file and attempt to make a unique molecule.
LSTSQ - To calculate structure factors and perform least squares from parameters in the CD file.
DISPOW - To calculate distances and angles and/or powder patterns.
ERRANL - To calculate error sums as functions of $\sin\theta/\lambda$, Fo, h,k and l.
TABLES - To produce structure factor and parameter listings for publications.
UTILTY - This consists of 11 utility-type routines to check, list or modify any .CD or .RE file.

Input data is derived from the NRC PDP-8e diffractometer control routine, though it is fairly easy to accept other forms of input data.

18.2-01 A Knowledge-based System for Interpreting Protein Electron Density Maps

A. Terry, R. Engelmore, Computer Science Department, Stanford University; S. Freer, Chemistry Department, UCSD; C. K. Johnson, Chemistry Division, ORNL

The partially realized goal for the computer program CRYSALIS is the derivation of protein structure models suitable for refinement, given 2.0-2.5 Ang. resolution density maps and amino acid sequences. The density map is represented by a graph derived through skeletonization and partitioning algorithms. The protein model under construction is represented by a node-link network called the hypothesis which is organized into three levels of detail: atoms, chemical groups, and secondary structures. The knowledge base pertaining to the problem is encoded into rule sets of situation-action rules where the "action" portion is executed only if the "situation" conditions are met. The top control level of the system selects the current operational strategy by using a set of strategy rules to evaluate the opportunities. A set of task rules utilizing the strategy selected examines the pending "events" which denote recent changes to the hypothesis. After the task rules have focused on an event, they select a third rule set which can make a new change to the hypothesis based on the event selected and the related status of the hypothesis. This change causes a new event and leads to further analysis by the task rule set. The control structure and overall design of CRYSALIS is based on recent Artificial Intelligence research in the design of knowledge-based systems. The programming languages are InterLISP and FORTRAN. An example of the system's performance is presented to show how the various kinds of knowledge are used in protein structure inference. Research at Stanford and at UCSD is supported by NSF; research at ORNL is supported by DOE.