energy barriers between minima, and subsequently attenuate the momenta again. In this way a number of minima may be found (Levitt & Warshel, 1975). It is equivalent to initializing a potential-energy minimization from a number of different conformations but it has the property that the minima so found are separated by energy barriers for which an upper limit is known so that the possibility exists of exploring transition pathways.

A second approach (Purisima & Scheraga, 1986) is relatively new. If the objective function to be minimized can be expressed in terms of interatomic distances, and if each atom is given coordinates in a space of $n - 1$ dimensions for n atoms, then a starting structure may be postulated for which the interatomic distances *all* take their ideal values and the objective function is then necessarily at an absolute minimum. This multidimensional structure is then projected into a space of fewer dimensions, within which it is again optimized with respect to the objective function. The dimensionality of the model is thus progressively reduced until a three-dimensional model is attained at a low energy. This means that the minimum so attained in three dimensions is approached *from beneath*, having previously possessed a lower value in a higher-dimensional space. This, in itself, does not guarantee that the three-dimensional minimum-energy structure so found is at the global minimum, but it is not affected by energy barriers between minima in the same way, and it does appear to reach very low minima, and frequently the global one. Because it is formulated entirely in terms of interatomic distances it offers great promise for modelling molecules on the basis of data from nuclear magnetic resonance.

3.3.3. Implementations

In this section the salient characteristics of a number of systems are described. Regrettably, it cannot claim to be a complete guide to all existing systems, but it probably describes a fairly representative sample. Some of these systems have arisen in academia and these are freely described. Some have arisen in or been adopted by companies which now market them, and these are described by reference to the original publications. Other marketed systems for which originators' published descriptions have not been found are not described. Yet other systems have been developed, for example, by companies within the chemical and pharmaceutical industries for their own use, and these have generally not been described in what follows since it is assumed that they are not generally available, even where published descriptions exist.

Software concerned especially with molecular dynamics has not been included unless it also provides static modelling capability, since this is a rapidly growing field and it has been considered to be beyond the intended scope of this chapter. Systems for which outline descriptions have already been given (Levitt & Lifson, 1969; Levitt, 1974; Diamond, 1966; Warme *et al.*, 1972; Dodson *et al.*, 1976; Hermans & McQueen, 1974) are not discussed further.

For some of the earliest work Levinthal (1966) still makes interesting reading and Feldmann (1976) is still an excellent review of the technical issues involved. The issues have not changed, the algorithms there described are still valuable, only the manner of their implementation has moved on as hardware has developed. A further review of the computer generation of illustrations has been given by Johnson (1980). Excellent bibliographies relating to these sections have been given by Morffew (1983, 1984), which together contain over 250 references including their titles.

The following material is divided into three sections. The first is concerned primarily with display rather than modelling though some of these systems can modify a model, the second is concerned with molecular modelling with reference to electron density and can develop a model *ab initio*, and the third is concerned with modelling with reference to other criteria.

Where software names are known to be acronyms constructed from initial letters, or where the original authors have used capitals, the names are capitalized here. Otherwise names are lower case with an initial capital.

While it is recognised that many of the systems here described are now of mainly historical interest, most have been retained for the second edition, some have been updated and some new paragraphs have been added.

3.3.3.1. *Systems for the display and modification of retrieved data*

One of the earliest systems designed for information retrieval and display was that described by Meyer (1970, 1971) which used television raster technology and enabled the contents of the Brookhaven Protein Data Bank (Meyer, 1974; Bernstein *et al.*, 1977) to be studied visually by remote users. It also enabled a rigid two-ring molecule to be solved from packing considerations alone (Hass *et al.*, 1975; Willoughby *et al.*, 1974). Frames for display were written digitally on a disk and the display rate was synchronized to the disk rotation. With the reduction in the cost of core storage, contemporary systems use large frame buffer memories thus avoiding synchronization problems and permitting much richer detail than was possible in 1970. A majority of the systems in this section use raster techniques which preclude real-time rotation except for relatively simple drawings, though *GRAMPS* is an exception (O'Donnell & Olson, 1981; Olson, 1982) (Section 3.3.3.1.4).

3.3.3.1.1. *ORTEP*

This program, the Oak Ridge Thermal Ellipsoid Program, due to Johnson (1970, 1976) was developed originally for the preparation of line drawings on paper though versions have since been developed to suit raster devices with interactive capability.

The program draws molecules in correct perspective with each atom represented by an ellipsoid which is the equi-probability surface for the atomic centre, as determined by anisotropic temperature factor refinement, the principal axes of which are displayed. Bonds are represented by cylindrical rods connecting the atoms which in the drawing are tapered by the perspective.

In line-drawing versions the problem of hidden-line suppression is solved analytically, whereas the later versions for raster devices draw the furthest elements of the picture first and either overwrite these with nearer features of the scene if area painting is being done or use the nearer features as erase templates if line drawings are being made.

3.3.3.1.2. *Feldmann's system*

R. J. Feldmann and co-workers (Feldmann, 1983) at the National Institutes of Health, Bethesda, Maryland, USA, were among the first to develop a suite of programs to display molecular structure using colour raster-graphics techniques. Their system draws with coloured shaded spheres, usually with one sphere to represent each atom, but alternatively the spheres may represent larger moieties like amino acids or whole proteins if lower-resolution representations are required. These workers have made very effective use of colour. Conventionally, oxygens have been modelled in red, but this system allows charged oxygens to be red and uncharged ones to be pink, with a similar treatment in blue for charged and uncharged nitrogens. By such means they have been able to give immediacy to the hydrophobic and electrostatic properties of molecular surfaces, and have used these characteristics effectively in studies of the binding possibilities of benzamidine derivatives to trypsin (Feldmann *et al.*, 1978).

The algorithm developed by Porter (1978) for shading spheres to be darkened near their peripheries also computes the proper appearance of the line of intersection of two spheres wherever interpenetration occurs, in contrast to some simpler systems which draw a complete disc for whichever sphere is forward of the other. Provided that all opaque spheres are drawn first, the system is also capable of representing transparent spheres by darkening the colour of the existing background inside, and especially near the edge of, discs representing transparent foreground spheres.

Other systems that produce space-filling pictures of a similar general character have been produced by Motherwell (1978), by Sundaram & Radhakrishnan (1979) and by Lesk (next section).

3.3.3.1.3. *Lesk & Hardman software*

The complexity of macromolecules is a formidable obstacle to perceiving the basic features of their construction and the stylized drawings produced by this software following the artistry of Richardson (1977, 1981, 1985) enables the internal organization of such molecules to be appreciated readily. The software is capable of mixing several styles of representation, among them the Richardson style of cylinders for α-helices, arrows for β-strands and ribbons for less-organized regions, or the creased-ribbon technique for the whole chain, or a ball-and-stick representation of atoms and bonds, or space-filling spheres. All these styles are available simultaneously in a single picture with depth cueing, colour and shading, and hidden-feature suppression as appropriate. It is also able to show a stylized drawing of a complete molecule together with a magnified part of it in a more detailed style. See Lesk & Hardman (1982, 1985).

3.3.3.1.4. *GRAMPS*

This system, due to O'Donnell & Olson (O'Donnell & Olson, 1981; Olson, 1982) provides a high-level graphics language and its associated interpretive software. It provides a general means of defining objects, drawable by line drawings, in such a way that these may be logically connected in groups or trees using a simple command language. Such a system may, for example, define a subunit protein of an icosahedral virus and define icosahedral symmetry, in such a way that modification of one subunit is expressed simultaneously in all subunits whilst preserving the symmetry, and simultaneously allowing the entire virus particle to be rotated. Such logical and functional relationships are established by the user through the medium of the *GRAMPS* language at run time, and a great diversity of such relationships may be created. The system is thus not limited to any particular type of structure, such as linear polymers, and has proved extremely effective as a means of providing animation for the production of cine film depicting viral and other structures. *GRAMPS* runs on all Silicon Graphics workstations under IRIX 4.0 or above.

3.3.3.1.5. *Takenaka & Sasada's system*

Takenaka & Sasada (1980) have described a system for the manipulation and display of molecular structures, including packing environments in the crystal, using a minicomputer loosely coupled to a mainframe. Their system is also capable of model building by the addition of groups of one or more atoms with a facility for monitoring interaction distances while doing so.

3.3.3.1.6. *MIDAS*

This system, due to Langridge and co-workers (Gallo *et al.*, 1983; Ferrin *et al.*, 1984) is primarily concerned with the display of existing structures rather than with the establishment of new ones, but it may modify such structures by bond rotations under manual control. It is of particular value in the study of molecular interactions since two or more molecules may be manipulated simultaneously and independently. Visual docking of molecules is greatly facilitated by the display of van der Waals surfaces, which may be computed in real time so that the turning of a bond in the underlying structure does not tear the surface (Bash *et al.*, 1983).

3.3.3.1.7. *Insight*

This system, originally due to Dayringer *et al.* (1986), has a functionality similar to *MIDAS*. It has been replaced by *Insight* II (current version 2.3.5). It appears to be well suited to the study of intermolecular relationships in docking and in structural comparisons, and it is able to make modifications to structures. Objects for display may be molecular or non-molecular, the former having an atomic substructure and the latter consisting of a vector list which may not be subdivided into referrable components. Map fitting with the current version has been reported.

3.3.3.1.8. *PLUTO*

PLUTO was developed by Motherwell (1978) at the Cambridge Crystallographic Data Centre (CCDC) for the display of molecular structures and crystal-packing diagrams, including an option for space-filling model style with shadowing. The emphasis was on a free format command and data structure, and the ability to produce ball-and-spoke drawings with line shadowing suitable for reproduction in journal publication. Many variant versions have been produced, with essentially the 1978 functionality, its popularity deriving from its ease of use and the provision of default options for establishing connectivity using standard bonding radii. It was distributed as part of the CCDC software associated with the Cambridge Structural Database, with an interface for reading entries from the database.

In 1993 Motherwell and others at the CCDC added an interactive menu and introduced colour and PostScript output. New features were introduced to allow interactive examination of intermolecular contacts, particularly hydrogen-bonded networks, and sections through packing diagrams (Cambridge Structural Database, 1994).

3.3.3.1.9. *MDKINO*

This system, due to Swanson *et al.* (1989), provides for the extraction and visualization of selected regions from molecular-dynamics simulations. It permits stereo viewing, interactive geometric interrogation and both forwards and backwards display of motion.

3.3.3.2. *Molecular-modelling systems based on electron density*

Systems described in this section require real-time rotation of complicated transparent scenes and all used vector-graphics technology in their original implementations for that reason, though many are now available for raster machines. In every case the graphics are the means of communication between the user and software possessing high functionality, capable of building a representation of a molecule *ab initio* and to alter it, change its shape and position it optimally in relation to an electron-density map, with due attention being paid to stereochemical considerations, by one or more of several approaches.

3.3.3.2.1. *CHEMGRAF*

Katz & Levinthal (1972) have developed a powerful modelling and display system for macromolecules known as *CHEMGRAF*. This system permits the definition of many atom types which includes bonding specifications, so that, for example, four types of carbon atom are included in the basic list and others may be added.

A molecular fragment with an unsatisfied valency (by which it might later be attached to another such group) would have that feature represented by a 'vanishing virtual' atom which removes the need for any organizational distinction between such fragments and complete molecules. Fragments, such as amino-acid residues, may be assembled from atoms, and molecules may be assembled from atoms and/or from such fragments invoked by name, by the superposition and elimination of the relevant vanishing virtuals. The assembly process includes the development of a connectivity tree for the molecule and provision is made for the 'turning' or reconstruction of such trees if the combination of such fragments redefines the root atom of one or more of the fragments. The system also provides for ring closure. Model building initially uses fixed bond lengths and angles, varying only the dihedral angles in single bonds, but has a library of preformed rings which could not otherwise be modelled on the simple basis. The results of such modelling may then be subjected to an energy-minimization routine using a steepest-descent method in the space of the dihedral angles and referring to the Lennard–Jones potential for non-bonded atom pairs. Atoms are first sorted into contiguous cubes so that all neighbours of any atom may be found by searching not more than 27 cubes.

The system is also capable of modelling by reference to electron density either by the translation and rotation of molecular fragments and the rotation of rotatable bonds within them or by the automatic linking of peaks in an electron-density map which are separated by less than, say, 1.8 Å, which is an important aid to interpretation when the resolution is sufficiently high.

3.3.3.2.2. *GRIP*

This system, developed by Professor F. P. Brooks, Dr W. V. Wright and associates at the University of North Carolina, Chapel Hill, NC, USA, was designed for biopolymers and was the first to enable a protein electron-density map to be interpreted *ab initio* without the aid of mechanical models (Tsernoglou *et al.*, 1977). Girling *et al.* (1980) give a more recent example of its use.

In its 1975 version *GRIP* is a three-machine system. Centrally there is a minicomputer which receives inputs from the user and controls a vector display with high-speed matrix-multiplication capability. The third machine is a mainframe computer with high-speed communication with the minicomputer.

The system develops a polymer chain from a library of monomers and manipulates it through bond rotations or free rotations of fragments explicitly specified by the user, with the aid of dials which may be coupled to bonds for the purpose. Bond rotations in the main chain either rotate part of the molecule relative to the remainder, which may have undesirable long-range effects, or the scope of the rotation is artificially limited with consequential discontinuities arising in the chain. Such discontinuities are removed or alleviated by the mainframe computer using the method of Hermans & McQueen (1974), which treats atomic position vectors, rather than bond rotations, as independent variables.

The system made pioneering use of a two-axis joystick to control orientation and a three-translation joystick to control position.

3.3.3.2.3. *Barry, Denson & North's systems*

These systems (Barry & North, 1971; North *et al.*, 1981; North, 1982) are examples of pioneering work done with minicomputers before purpose-built graphics installations became widespread; examples of their use are given by Ford *et al.* (1974), Potterton *et al.* (1983) and Dodson *et al.* (1982). They have the ability to develop a polymer chain in sections of several residues, each of which may subsequently be adjusted to remove any misfit errors where the sections overlap. Manipulations are by rotation and translation of sections and by bond rotations within sections. These movements are directly controlled by the user, who may simultaneously observe on the screen the agreement with electron density, or calculated estimates of potential or interaction energy, or a volume integral of the product of observed and model densities, or predicted shifts of proton magnetic resonance spectra. Thus models which are optimal by various criteria may be constructed, but there is no optimizer directly controlling the rotational adjustments which are determined by the user.

One of the earliest applications of them (Beddell, 1970) was in the fitting of substrate molecules to the active site of lysozyme using difference electron densities; however, the systems also permitted the enzyme–substrate interaction to be studied simultaneously and to be taken into account in adjusting the model.

3.3.3.2.4. *MMS-X*

The Molecular Modelling System-X (*MMS-X*) is a system of purpose-built hardware developed by Barry, Marshall and others at Washington University, St Louis. Associated with it are several sets of software. The St Louis software consists of a suite of programs rather than one large one and provides for the construction of a polymer chain in helical segments which may be adjusted bodily to fit the electron density, and internally also if the map requires this too. Non-helical segments are built helically initially and unwound by user-controlled rotations in single bonds. The fitting is done to visual criteria. An example of the use of this system is given by Lederer *et al.* (1981).

Miller *et al.* (1981) have described an alternative software system for the same equipment. Functions invoked from a keyboard allow dials to be coupled to dihedral angles in the structure. Their system communicates with a mainframe computer which can deliver small blocks of electron density to be contoured and stored locally by the graphics system; this provides freedom of choice of contour level at run time. An example of the use of this system is given by Abad-Zapatero *et al.* (1980).

3.3.3.2.5. *Texas A&M University system*

This system (Morimoto & Meyer, 1976), a development of Meyer's earlier system (Section 3.3.3.1), uses vector-graphics technology and a minicomputer and is free of the timing restrictions of the earlier system. The system allows control dials to be dynamically coupled by software to rotations or translations of parts of the structure, thus permitting the re-shaping or re-positioning of the model to suit an electron-density map which may be contoured and managed by the minicomputer host. The system may be used to impose idealized geometry, such as planar peptides in proteins, or it may work with non-idealized coordinates.

The system was successfully applied to the structures of rubredoxin and the extracellular nuclease of *Staphylococcus aureus* (Collins *et al.*, 1975) and to binding studies of sulfonamides to carbonic anhydrase (Vedani & Meyer, 1984). In addition, two of the first proteins to be constructed without the aid of a 'Richards' Box' were modelled on this system: monoclinic lysozyme in 1976 (Hogle *et al.*, 1981) and arabinose binding protein in 1978 (Gilliland & Quiocho, 1981).

3.3.3.2.6. *Bilder*

This system (Diamond, 1980*a*,*c*, 1982*b*) runs on a minicomputer independent of any mainframe. It builds a polymer chain from a library of residues and adapts it by internal rotations and overall positioning in much the same way as previous systems described in this section. Like them, it can provide user-controlled bond rotations, but its distinctive feature is that it has an optimizer within the minicomputer which will determine optimal combinations of bond rotations needed to meet the user's declared

objectives. Such objectives are normally target positions for atoms set by the user by visual reference to the density, using the method of Section 3.3.1.3.9, but they may include target values for angles. These latter may either declare a required shape that is to take precedence over positional requirements, which are then achieved as closely as the declared shape allows, or they may be in least-squares competition with the positional requests. The optimizer also recognizes the constraints imposed by chain continuity and enables an internal section of the main chain to be modified without breaking its connection to the rest of the molecule. Similar techniques also allow ring systems to adopt various conformations, by bond rotation, without breaking the ring, simultaneously permitting the ring to have target positions. The optimizer is unperturbed by under-determined situations, providing a minimum-disturbance result in such cases. All these properties of the optimizer are generated without recourse to any 'special cases' by a generalization of the subspace section technique which was used to maintain chain continuity in a 'real-space-refinement' program (Diamond, 1971). This is based entirely on the rank of the normal matrix that arises during optimization, which may serve to satisfy a constraint such as chain continuity or ring closure and simultaneously to establish what degrees of freedom remain to be controlled by other criteria. In *Bilder* this is achieved without establishing eigenvalues or eigenvectors. The method is described in outline in Section 3.3.2.2.1 and in detail by Diamond (1980*a*,*b*).

The angular variables used normally comprise all single bonds but may include others, such as the peptide bond with or without a target of 180°. Thus this bond may be completely rigid, elastic, or completely free. Any interbond angles may also be parameterized but at some cost in storage. The normal mode of working is to develop a single chain for the entire length of the molecule, but if cumulative error makes fitting difficult a fresh chain may be started at any stage. *Bilder* may itself reconnect such chains at a later stage.

Construction and manipulation operates on a few residues at a time within the context of a polymer chain, but any or all of the rest of the molecule, or other molecules, may be displayed simultaneously.

Contouring is done in advance to produce a directoried file of contoured bricks of space, each brick containing up to 20 independently switchable elements which need not all be from the same map. Choice of contour level and displayed volume is thus instantaneous within the choices prepared.

The system is menu driven from a tablet, only file assignments and the like requiring the keyboard, and it offers dynamic parallax as an aid to 3D perception (Diamond *et al.*, 1982). Bloomer *et al.* (1978), Phillips (1980), and Evans *et al.* (1981) give examples of its use.

3.3.3.2.7. *Frodo*

This system, due to Jones (Jones, 1978, 1982, 1985; Jones & Liljas, 1984), in its original implementation was a three-machine system comprising graphics display, minicomputer and mainframe, though more recent implementations combine the last two functions in a 'midi'. Its capabilities are similar to those of *Bilder* described above, but its approach to stereochemical questions is very different. Where *Bilder* does not allow an atom to be moved out of context (unless it comprises a 'chain' of one atom) *Frodo* will permit an atom or group belonging to a chain to be moved independently of the other members of the chain and then offers regularization procedures based on the method of Hermans & McQueen (1974) to regain good stereochemistry. During this regularization, selected atoms may be fixed, remaining atoms then adjusting to these. A peptide, for example, may be inverted by moving the carbonyl oxygen across the peptide and fixing it, relying on the remaining atoms to rearrange themselves. (*Bilder* would do the equivalent operation by cutting the chain nearby, turning the peptide explicitly, reconnecting the chain and optimizing to regain chain continuity.) The *Frodo* approach is easy to use especially when large displacements of an existing structure are called for, but requires that ideal values be specified for all bond lengths, angles and fixed dihedrals since the system may need to regain such values in a distorted situation. *Bilder*, in contrast, never changes such features and so need not know their ideal values.

Frodo may work either with consecutive residues of a polymer chain, useful for initial building, or with a volume centred on a chosen position, which is ideal for adjusting interacting side chains which are close in space but remote in sequence.

In recent implementations *Frodo* can handle maps both in density grid form and in contour form and permits on-line contouring. It has also been developed (Jones & Liljas, 1984) to allow the automatic adjustment of the position and orientation of small rigid groupings by direct reference to electron density in the manner of Diamond (1971) but without the maintenance of chain continuity, which is subsequently reintroduced by regularization.

Horjales and Cambillau (Cambillau & Horjales, 1987; Cambillau *et al.*, 1984) have also provided a development of *Frodo* which allows the optimization of the interaction of a ligand and a substrate with both molecules being treated as flexible.

3.3.3.2.8. *Guide*

Brandenburg *et al.* (1981) have described a system which enables representations of macromolecules to be modified with reference to electron density. Such modifications include rotation about single bonds under manual control, or the movement with six degrees of freedom, also under manual control, of any part or parts of the molecule relative to the remainder. The latter operation may necessitate subsequent regularization of the structure if the moved and unmoved parts are chemically connected, and this is done as a separate operation on a different machine. The system also has the capability of displaying several molecules and of manually superimposing these on each other for comparison purposes.

3.3.3.2.9. *HYDRA*

This program, due to Hubbard (1985) (and, more recently, to Molecular Simulations) has several functional parts, referred to as 'heads', which all use the same data structure. The addition of further heads may be accomplished, knowing the data structure, without the need to know anything of the internal workings of existing heads.

The program contains extensive features for the display, analysis and modelling of molecular structure with particular emphasis on proteins. Display options include dotted surfaces, molecular skeletons, protein cartoons and a variety of van der Waals, ball-and-stick, and other raster-graphics display techniques such as ray tracing and shaded molecular surfaces. Protein analysis features include the analysis of hydrogen bonding, and of secondary and domain structure, as well as computational assessment of deviations from accepted protein structural characteristics such as abnormal main-chain or side-chain conformations and solvent exposure of hydrophobic amino acids. A full set of protein modelling facilities are provided including homology modelling and the 'docking' of substrate molecules. The program contains extensive tools for interactive modelling of structures from NMR or X-ray crystallographic data, and provides interfaces to molecular-mechanics and dynamics calculations. There are also database searching facilities to analyse and compare features of protein structure, and it is well suited to the making of cine films.

3.3.3.2.10. *O*

Jones *et al.* (1991) have developed a modelling system for proteins with a radically different approach to any of the foregoing, in that they begin by reducing the available electron-density map to a skeletal representation (Greer, 1974; Williams, 1982) which consists of a line running through the density close to its maximal values, this being the basis of a chain trace. Provisional α-carbon positions are also estimated at this stage. A database of known structures is then scanned for pentapeptides which may be superimposed on five successive positions in the chain trace, the best fit so found being taken to provide coordinates for the three central residues of the developing model. The process advances by three residues at each step, the first and last residues of the pentapeptide being used only to ensure that the central residues are built in a manner compatible with what precedes and follows.

The process ensures that conformations so built are free from improbable conformations, and the whole forms an adequate starting structure for molecular-dynamics procedures, even though some imperfect geometry is to be expected where each tripeptide joins the next.

3.3.3.3. *Molecular-modelling systems based on other criteria*

Systems described within this section mostly have some form of energy minimization as their objective but some are purely geometrical. The optimization of molecules through empirical force fields has been reviewed by Allinger (1976), Burkert & Allinger (1982) and Boyd & Lipkowitz (1982). Some of these systems are in the academic domain, others are commercial. Most have capabilities exceeding the features referred to here and, of necessity, the list cannot be complete. No attempt at comparative evaluations is attempted or implied.

3.3.3.3.1. *Molbuild, Rings, PRXBLD and MM2/MMP2*

Liljefors (1983) has described a system for constructing representations of organic molecules. The system develops the molecule with plausible geometry and satisfied valencies at all stages of the development with explicit recognition of lone pairs and the various possible hybridization states. Growth is generally by substitution in which a substituent and the atom it is to replace are both nominated from the screen. The bond which is reconstructed in a substitution is generally a single bond. Double and triple bonds are introduced by the substitution of moieties containing them. Atom types may be changed after incorporation in the growing molecule, so that although the menu of substituents includes —CH_3 but not —NH_2 the latter may be obtained by incorporating —CH_3, then changing C to N and one of the hydrogens to a lone pair. Facilities are also provided for cyclization and acyclization.

van der Lieth *et al.* (1984) have described an extension to this that is specialized to the construction of fused-ring systems. It permits the joining of rings by fusion of a bond, in which two adjacent atoms in one ring are superposed on two in another. It also permits the construction of spiro links in which one atom is common to two rings, or the construction of bridges, or the polymerization of ring systems to form, for example, oligosaccharides. Again the satisfaction of valencies is maintained during building and the geometry of the result is governed by superposition of relevant atoms in the moieties involved.

PRXBLD is a molecular model-building program which accepts two-dimensional molecular drawings in a manner similar to *Script* (Section 3.3.3.3.2) and constructs approximate three-dimensional coordinates from these. It is the model-building component of *SECS* (Simulation and Evaluation of Chemical Synthesis) (Wipke *et al.*, 1977; Wipke & Dyott, 1974; Wipke, 1974). See also Anderson (1984).

All three of these programs produce output which is acceptable as input to *MM*2(82)/*MMP*2 which are developments of Allinger's geometrical optimization based on molecular mechanics (Allinger, 1976).

3.3.3.3.2. *Script*

This system, described by Cohen *et al.* (1981), is specialized for fused-ring systems, especially steroids, but is not limited to these classes. The system allows the user to draw on the screen (with a light pen or equivalent) a two-dimensional representation of a molecule using single lines for single bonds, double lines for double bonds, and wedges to indicate out-of-plane substituents. The software can then enumerate the possible distinct conformers, each of which is expected to be near an energy minimum on the conformational potential surface. Each conformer may then be annealed to reach an energy minimum using an energy estimate based on bond lengths, bond angles, torsion angles and van der Waals, electrostatic and hydrogen-bonding terms. An example is given of the identification of an unusual conformer as the most stable one from twelve possibilities for a four-ring system.

The program is a development of similar work by Cohen (1971) in which the molecule was defined in terms of a tree structure and an optimizer based on search techniques rather than gradient vectors was used. The method included van der Waals terms and hence estimated energy differences between stereoisomers in condensed ring systems arising from steric hindrance.

3.3.3.3.3. *CHARMM*

This system, due to Brooks *et al.* (1983), is primarily concerned with molecular dynamics but it includes the capability of model-building proteins and nucleic acids from sequence information and values of internal coordinates (bond lengths, bond angles and dihedral angles). The resulting structure (or a given structure) may then be optimized by minimizing an empirical energy function which may include electrostatic and hydrogen-bonding terms as well as the usual van der Waals energy and a Hookean treatment of the covalent skeleton. Hydrogen atoms need not be handled explicitly, groups such as —CH_2— being treated as single pseudo atoms, and this may be advisable for large structures. For small or medium proteins hydrogens may be treated explicitly and their initial positions may be determined by *CHARMM* if they are not otherwise known.

3.3.3.3.4. *Commercial systems*

A number of very powerful molecular-modelling systems are now available commercially and we mention a few of these here. Typically, each consists of a suite of programs sharing a common data structure so that the components of a system may be acquired selectively.

The *Chem-X* system, from Chemical Design Ltd, enables models to be developed from sketch-pad input, provides for their geometrical optimization and interfaces the result to *Gaussian*80 for quantum-mechanical calculations.

MACCS, from Molecular Design Ltd, and related software (Allinger, 1976; Wipke *et al.*, 1977; Potenzone *et al.*, 1977) has similar features and also has extensive database-maintenance facilities including data on chemical reactions.

Sybyl, from Tripos Associates (van Opdenbosch *et al.*, 1985), also builds from sketches with a standard fragment library, and provides interfaces to quantum-mechanical routines, to various databases and to *MACCS*.

Insight II (Section 3.3.3.1.7) is available from Biosym and *GRAMPS* (Section 3.3.3.1.4) is available from T. J. O'Donnell Associates.