The Structure of Atomic and Molecular Clusters, Optimised using Classical Potentials

M. M. Ali

School of Computational and Applied Mathematics
University of the Witwatersrand
Johannesburg, South Africa
mali@cam.wits.ac.za

Roger Smith and S. Hobday
Mathematical Sciences,
Loughborough University,
Loughborough LE11 3TU, UK
R.Smith@lboro.ac.uk

May 15, 2006

Abstract

The problem of the determination of the minimum energy configuration of an arrangement of N point particles under the interaction of their interatomic forces is discussed. The interatomic forces are described by classical many body potentials. Different optimisation methods are considered, multi level single link, topographical differential evolution and a genetic algorithm but it is shown that genetic algorithms combined with an efficient local optimisation method is especially quick and reliable for this task. In addition to comparing some different optimisation methods, the structures of clusters of atoms described by interatomic potential functions containing up to a few hundred atoms are calculated including some with some special symmetries. A number of applications are given including covalent carbon and silicon clusters, close-packed structures such as argon and silver and the two-component carbon-hydrogen system.

Keywords: Optimisation, Atomic Clusters, Potential Energy Surfaces, Many-body Potentials.

1 Introduction

This paper addresses the problem of determining the lowest energy configuration of a system of particles. This is an old problem related to that of determining the optimal arrangement of close-packed spheres, which has attracted the attention of pure mathematicians for many years. The "orange box" arrangement is generally thought to be the best way of packing the most number of spheres together into the least possible volume but a formal proof of this has proved to be very difficult. In this arrangement, the first layer consists of spheres in a close-packed triangular

lattice. The second layer is formed by placing a sphere in the centre of the depressions left in the first layer, forming a lattice shifted with respect to the first layer. In the third layer the spheres can lie either directly above those in the first layer or in the unused depressions from either layer. In the first case this is known as hexagonal close packing (hcp) and in the second face-centred-cubic (fcc) because the spheres are arranged in a periodic cubic unit cell at the faces of a cube. In many materials, especially metals and frozen rare gases, the atoms also arrange themselves into this fcc structure. For other types of materials where directional bonding is important, such simple close-packed structures are not so common and complex geometrical arrangements occur depending precisely on the nature of the chemical interactions.

In this paper we address the problem from another angle. The spheres are now considered to be atoms that are regarded as point particles which interact through their interatomic forcefields. The forcefields are given as the starting point for the problem and are determined from inter-atomic potentials. The problem is therefore to find the optimal arrangement, i.e. the minimum potential energy, when N such particles interact. At the most sophisticated level, the forcefields are determined by the solution of Schrödinger's equation under various levels of approximation. This is a very complicated and time-consuming procedure which although becoming more standard can still take many days of computing time even to calculate the forces between 10's of atoms.

To overcome this problem, semi-empirical interatomic potential functions have been developed. They are of various forms and levels of detail depending on the type of material. For example, covalent materials such as silicon and carbon have strong directional bonds and can be modelled by so-called bond order potentials [1, 2, 3]; metals can be described by a model which describes the atomic nucleii as being embedded in a sea of electrons, the so-called embedded atom potentials [4, 5]. Some rare gases [6] can even be described by simple potentials that are pairwise additive such as the well-known Lennard-Jones potential. Multi-component materials can also be described. For more details on the forms of these types of potential function see the review article by Carlsson [7].

The energy minimisation task is an extremely challenging global optimisation problem for even moderately sized clusters, because the number of variables as well as the number of local minima are very high. The problem can therefore be used as a sample test problem for global optimisation. In order to calculate the minimum potential energy for small number of particles using some of the potential functions described below, we first attempted to use several recent stochastic global optimisation techniques. They are aspiration based simulated annealing [8], multi-level single linkage [9], topographical multi-level single linkage [10] and controlled random search [11]. Our numerical experiments with these algorithms suggested that they are successful and robust for problems with a small number of dimensions as well as problems with a small number of local minima. For problems containing up to five particles most of these algorithms were robust in locating the best minimum. However, as the number of particles increases their robustness diminishes. We will present the results of these algorithms in the comparison section. The results obtained by the algorithms suggested that for the energy minimisation problems with high dimensions, special types of algorithms are required. Therefore in this paper we will concentrate on two recent global optimisation techniques designed for energy minimisation problems. These are the topographical differential evolution [12] and the genetic algorithm [13]. There are several genetic algorithms have been designed for chemical cluster optimisation previously. For more details of these see the review articles by Judson [14] and [15].

The most extensively tested energy minimisation problem is that using the Lennard-Jones potential. Extensive numerical studies have been carried out on this pair potential function, for

some recent studies see [16, 17, 18]. The number of local minima for this problem is believed to be exponential in the number of atoms or molecules. The Lennard-Jones (L-J) function is simple to state and easily programmable, yet challenging and complex in the behavior of its solutions.

2   Potential functions

In this section, we describe some well known potential functions that can be used as challenging test problems for global optimisation.

2.1   Lennard-Jones potential function

The Lennard-Jones problems assumes that the potential energy of a cluster of atoms is given by the sum of the pairwise interactions between atoms, with these interactions being Van der Waals forces given by the Lennard-Jones $6 - 12$ potential. That is, if we define the position of the molecular cluster of N_a atoms by

$$x = (x_1, x_2, \cdots, x_{N_a})$$

where x_i is a three-dimensional vector denoting the coordinates of the ith atom, then the potential energy function in non-dimensional form [6] is

$$E(x) = 4\epsilon \sum_{i=1}^{N_a} \sum_{j=1}^{i-1} \left[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^6 \right] \tag{1}$$

where $r_{ij} = |x_i - x_j|$ is the Euclidean distance between x_i and x_j, σ and ϵ are constants that depends on the type of atom.

2.2   The embedded atom potential function for metals

In the framework of the embedded atom formalism, the energy of an assembly of N_a atoms is given by

$$E(x) = \frac{1}{2} \sum_{i \neq j}^{N_a} V_{ij}(r_{ij}) - \sum_{i=1}^{N_a} (\sum_{j=1}^{N_a} \Phi_{ij}(r_{ij}))^{\frac{1}{2}}$$

where V_{ij} is the pairwise repulsive part of the potential energy and the term containing $\Phi(r_{ij})$ in the many body isotropic, cohesive term. $\Phi(r_{ij})$ itself is a pairwise additive function similar to the V_{ij} term and the Lennard-Jones potential described above. This form of potential energy function has been used to describe both fcc, hcp and also body-centred-cubic materials, depending on the form of the functions V and Φ and the parameters that define them. See for example [5] which gives these functions and parameters for some fcc metals.

2.3   The Tersoff potential for carbon and silicon

The potential function due to Tersoff was originally designed for atomic interaction of silicon atoms where there is strong covalent (i.e directional) bonding. The many-body term is no longer isotropic as in the case of the embedded atom potential. Tersoff considered two different parameterisations of silicon which we will call, $Si(A)$ and $Si(B)$. These have slightly different

properties with regard to surfaces and bulk elasticity. There is also a parameterisation to describe carbon. The binding energy in the Tersoff formulation [2] is written as a sum over atomic sites in the form

$$E_i = \frac{1}{2} \sum_{j \neq i} f_c(r_{ij})(V_R(r_{ij}) - \beta_{ij} V_A(r_{ij})) , \quad \forall i \tag{2}$$

where r_{ij} is the distance between atoms i and j, V_R is a repulsive term, V_A is an attractive term, $f_c(r_{ij})$ is a switching function that reduces the potential smoothly to zero at large distances and β_{ij} is a many-body term that depends on the positions of atoms i and j and the neighbours of atom i. More details of each of these quantities can be found in [2]. The many-body term β_{ij} is based on atom i, so that in general $\beta_{ij} \neq \beta_{ji}$. The term β_{ij} is given by

$$\beta_{ij} = (1 + \gamma^{n_1} \xi_{ij}^{n_1})^{-1/2n_1} \tag{3}$$

where n_1 and γ are known fitted parameters [2]. The term ξ_{ij} for atoms i and j (i.e., for bond ij) is given by

$$\xi_{ij} = \sum_{k \neq i,j} f_c(r_{ik}) g(\theta_{ijk}) \exp\left(\lambda_3^3 (r_{ij} - r_{ik})^3\right) , \tag{4}$$

The term ξ_{ij} describes the contribution of the neighbours of the atom i, ξ_{ij} increases as the number of k atoms increases but the term β_{ij} decreases as ξ_{ij} increases. The exponential term in (4) is designed to reduce the contribution of bonds with length greater than r_{ij}, so that the distant neighbours of i have a reduced contribution to the bond order term. The term θ_{ijk} is the bond angle between bonds ij and ik and g is given by

$$g(\theta_{ijk}) = 1 + c^2/d^2 - c^2/[d^2 + (h - \cos\theta_{ijk})^2] . \tag{5}$$

The parameter h is the cosine of the optimum bond angle and c and d control the influence of bond angles on the many-body term. The quantities λ_3, c, d and h which appear in (4) and (5) are also known fitted parameters. The terms $V_R(r_{ij})$ and $V_A(r_{ij})$ are given by

$$\begin{aligned} V_R(r_{ij}) &= A e^{-\lambda_1 r_{ij}} &(6a) \\ V_A(r_{ij}) &= B e^{-\lambda_2 r_{ij}} &(6b) \end{aligned}$$

where A, B, λ_1 and λ_2 are given fitted parameters. The switching function $f_c(r_{ij})$ restricts the potential calculations to nearest neighbours only and ensures that atomic interactions decay smoothly to zero as the seperation distance increases from R_1 to R_2. It is given by

$$f_c(r_{ij}) = \begin{cases} 1 , & r_{ij} \leq R_1 \\ \frac{1}{2} - \frac{1}{2}\sin[\pi(r_{ij} - R)/(R_2 - R_1)] , & R_1 < r_{ij} < R_2 . \\ 0 , & r_{ij} \geq R_2, \end{cases} \tag{7}$$

where $R = 0.5(R_1 + R_2)$. The parameters for the different materials are given in the original references.

2.4 Brenner hydrocarbon potential

The binding energy for Brenner hydrocarbon potential [19, 20] is similar to Tersoff potential and is given by

$$E_i = \frac{1}{2} \sum_{j \neq i} f_c(r_{ij}) \left(V_R(r_{ij}) - \bar{\beta}_{ij} V_A(r_{ij})\right) . \quad \forall i \tag{8}$$

4

The subscriptions i and j refer to first neighbour atoms separated by the distance r_{ij}. Here we do not give all the detailed parameters and functions for the potential which can be found in the original papers. Instead we point out only the principle differences between this and the Tersoff potential. The attractive and repulsive terms in Brenner potential are given by

$$
\begin{aligned}
V_R(r_{ij}) &= \frac{D_o}{S-1} \exp\left(-\beta\sqrt{2S}(r_{ij}-R_e)\right) \qquad (9a)\\
V_A(r_{ij}) &= \frac{SD_o}{S-1} \exp\left(-\beta\sqrt{2/S}(r_{ij}-R_e)\right) \qquad (9b)
\end{aligned}
$$

where D_o is the energy at a separation R_e and these are known as dimer energy and the dimer separation respectively, and S and β are constants. The cut-off function $f_c(r_{ij})$ restricts the potential to nearest neighbour interactions as before.

Brenner's first correction is to take into account non-local effects which means quantitative information on the second neighbours of both atoms i and j and this is used to account for conjugated versus non-conjugated double bonds (see Figure 1).
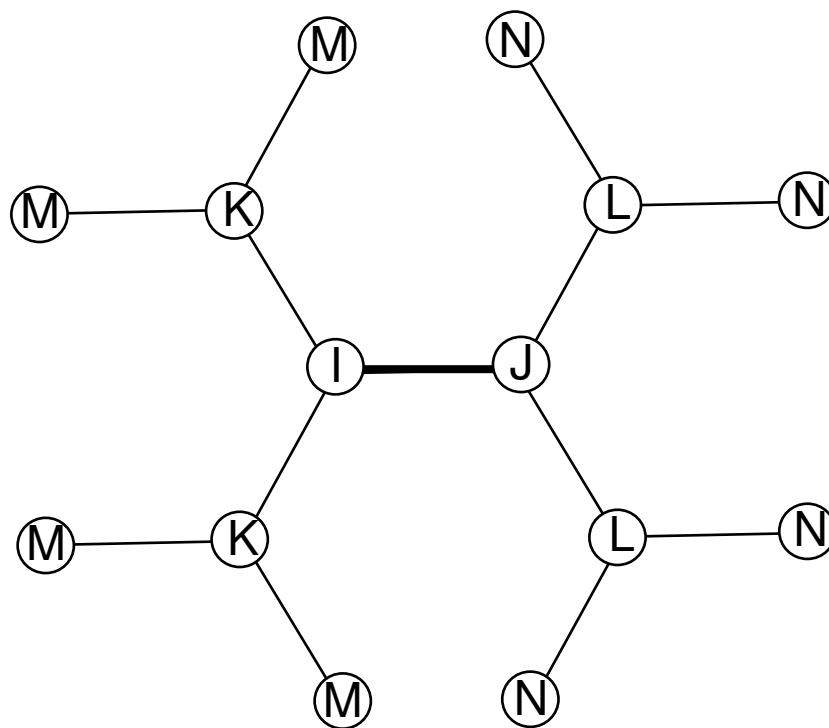

Figure 1: In the graphite (hexagonal lattice) structure, the energy of the central bond between atoms i and j is a function of the positions of i and j as well as their first neighbour atoms, excluding i and j (the four positions: two Ks and two Ls) and second neighbour atoms (the eight positions: four Ms and four Ns).

The bond order function $\bar{\beta}_{ij}$, given below, is the average of terms associated with atoms i and j in the bond ij plus a correction.

$$
\bar{\beta}_{ij} = \frac{1}{2}\left(\beta_{ij}+\beta_{ji}\right) + \frac{1}{2}F_{ij}\left(N_i^t, N_j^t, N_{ij}^{conj}\right) . \qquad (10)
$$

The average term here is to ensure that all first bond neighbours are accounted for; β_{ij} takes into account the first neighbours of atom j and β_{ji} accounts for first neighbours of atom i. The contribution of atom i to the bond order term is thus given by

$$\beta_{ij} = \left[1 + \xi_{ij} + H_{ij}\left(N_i^H, N_i^C\right)\right]^{-\delta_i} . \tag{11}$$

Thus the Brenner potential does take into account conjugation between atoms, a feature that was missing from the original Tersoff formulation.

3 Global optimisation algorithms

In this section we describe four algorithms that are suitable for use in solving the problem described above. They are multilevel single linkage (MSL) [9], topographical multilevel single linkage (TMSL) [10], topographical differential evolution (TDE) [12] and the genetic algorithm (GA) [13]. All are iteration (or generation) based algorithms. Of these algorithms, MSL and TMSL have a similar algorithmic structure. They generate random points from the search space per iteration. On the other hand, TDE and GA do not generate random points per iteration. They maintain a set (population set) of candidate solutions (points in the search space), the size of which does not vary with iteration. They are known as population based algorithms. Although they do not generate random points per iteration, they create new points, known as the children, using some genetic operators such as crossover and mutation. Newly created children in an iteration replace bad parents (bad candidate solutions) in the population set. Initial points in the population set are generated randomly within the search space, say Ω.

3.1 The MSL and TMSL algorithms

MSL and TMSL are stochastic algorithms in the sense that they have probabilistic convergence guarantee [9, 32]. The only difference between the two is that MSL uses a special clustering technique, called 'multilevel single linkage' and TMSL uses a 'topographical clustering technique', per iteration. Suffice it therefore to describe one algorithm, e.g. the MSL algorithm.

In the j-th iteration of the MSL algorithm, N new sample points are randomly generated from the search space, Ω. These points are then added to the the $(j-1)N$ points generated in all previous iterations, making the sample point of size jN at the j-th iteration. The sample size jN is then reduced by removing the $(1-\gamma_1)jN$ points with the highest function values. Typically, $\gamma_1 = 0.15$ is used. A number of clusters at the j-th iteration is then formed using $\gamma_1 jN$ points. A local search is then performed from a potential point of each cluster. The MSL algorithm stops when a convergence criterion is met. In the topographical clustering used by TMSL, the centre of a cluster is known as the 'graph minimum'. TMSL performs local searches from the graph minima of clusters.

Notice that clusters formed by the MSL and TMSL methods have essentially a different meaning from the cluster of atoms or particles. A cluster of atoms is a point in the search space Ω where the minimum is sort. On the other hand, the clustering involved in either of the algorithms is to group together a number of (reduced) sample points in the search space where each point in a group is a cluster of atoms. Therefore, if $x = (x_1, x_2, \cdots, x_{N_a})$ then $x \in \Omega \subset IR^{3N_a}$ and $x_i \in IR^3$.

3.2 Topographical differential evolution (TDE)

Unlike the GA, TDE uses two population sets. Initially, two sets S_1 and S_2, each containing k points, are generated in the following way; iteratively sample two points from the search region, Ω, the best point x^i going to S_1 and the other $x^{i'}$ to S_2. The process continues until each set has k points. The TDE procedure then gradually drives both the sets S_1 and S_2 towards the global minimiser through the repeated cycles of mutation, crossover, acceptance and replacement. In each cycle constituting a generation or iteration, k competitions are held to determine the members of S_1 and S_2 for the next generation. The i-th ($i = 1, 2, \cdots, k$) competition is held to replace x^i in S_1. Considering x^i as the target point, a trial point y^i is found from two points (parents), the point x^i, i.e., the target point and the point $\hat{x}^i$ determined by the mutation operation.

In its mutation phase TDE randomly selects three distinct points $x^{p(1)}$, $x^{p(2)}$ and $x^{p(3)}$ from the current set S_1. None of these points should coincide with the current target point x^i. The weighted difference of any two points is then added to the third point which can be mathematically described as :

$$\hat{x}^i = x^{p(1)} + F(x^{p(2)} - x^{p(3)}) \tag{12}$$

where $F > 0$ is a scaling factor, and $x^{p(1)}$ is known as the base vector. If the point $\hat{x}^i \notin \Omega$ then the mutation operation is repeated.

The trial point $y^i = (y_1^i, y_2^i, \cdots, y_{N_a}^i)$ is found from its parents x^i and $\hat{x}^i$ using the following crossover rule. N_a random numbers, R^j ($j = 1, 2, \cdots, N_a$), in (0,1) are generated repeatedly. If the number is greater than 0.5 then the corresponding component of y^i is taken from $x^i = (x_1^i, x_2^i, \cdots, x_{N_a}^i)$ else it is taken from $\hat{x}^i = (\hat{x}_1^i, \hat{x}_2^i, \cdots, \hat{x}_{N_a}^i)$. The j-component of y^i is therefore calculated as follows:

$$y_j^i = \begin{cases} \hat{x}_j^i & \text{if } R^j \leq 0.5 \\ x_j^i & \text{if } R^j \geq 0.5. \end{cases} \tag{13}$$

In the acceptance phase, if the trial point y^i, corresponding to the target x^i, does satisfy the criterion $f(y^i) < f(x^i)$ then the point y^i replaces x^i in S_1, if however, y^i does not satisfy the above criterion then it is not abandoned altogether, rather it competes with its corresponding target $x^{i'}$ in the set S_2. If $f(y^i) < f(x^{i'})$ then y^i replaces $x^{i'}$ in S_2. The potential points in S_2 then can be used for further exploration and exploitation.

As the search proceeds TDE introduces two measures to lessen the chance of missing the global minimiser in the driving process of S_1 and S_2. They are : (a) after each M iterations, determine the graph minima [1] using the N_g best points from S_1 and then perform a local search from each of the graph minima found, and (b) the replacement of the worst N_g points in S_1 with the best N_g points in S_2 immediate after the local searches have been performed. The graph minima used for local search are the centres of the topographical clusters.

The benefits of (a) are that a local search only starts from a potential point with low function value and these potential points are seperated by higher regions. Since the points in S_1 gradually shift their position these periodically scrutinised local searches will enhance the robustness of the TDE algorithm in locating the global minimum. The benefits of (b) are search diversification and exploitation. TDE repeatedly finds the graph minima locally using N_g best points with k_g nearest neighbours. The best minimum found in the local search phase is recorded and is

---

[1] A simplified description of how to find the graph minima is the following: For each point in the sample of size N_g ($N_g < k$) the k_g nearest neighbour points are determined ($k_g \ll N_g$). Those points for which all k_g near neighbours are inferior points, i.e., the function values are larger, are the graph minima.

further updated in the next phase of local search. If a consecutive number, say t, of local search phases does not produce any better minimum value than the previously found best minimum then the algorithm can be terminated. The step by step description of the new algorithm is as follows.

The TDE Algorithm

Step 1 Determine the initial sets $S = \left\{x^1, x^2, \cdots, x^k\right\}$ and $S_a = \left\{x^{1'}, x^{2'}, \cdots, x^{k'}\right\}$ with points sampled randomly in Ω. Initialize the generation counter i_g and the local phase counter t to zero.

Step 2 If the stopping condition, $t \geq t_{max}$, is satisfied then stop.

Step 3 For each $x^i \in S_1$, determine y^i by the following two operations:

- Mutation : Randomly select three points from S_1 except x^i, the running target and find the second parent $\hat{x}^i$ by the mutation rule (12).
- Crossover : Calculate the trial vector y^i corresponding to the target x^i from x^i and $\hat{x}^i$ using the crossover rule (13).

Step 4 Update both the sets S_1 and S_2 for the next generation using the acceptance rule: replace each $x^i \in S_1$ with y^i if $f(y^i) < f(x^i)$ otherwise replace $x^{i'} \in S_2$ with y^i if $f(y^i) < f(x^{i'})$. Set $i_g := i_g + 1$. If $i_g \equiv 0 \,(\mathrm{mod}\, M)$ then go to Step 5, otherwise go to Step 2.

Step 5 Find the graph minima of the function, $f(x)$, using the best N_g points in S_1 and perform a local search starting from each graph minimum. Keep a record of the very best minimum found so far, replace the worst N_g points in S_1 with the best N_g in S_2. If the current phase of local minimisation produces a better minimum than the current best minimum then set $t = 0$ otherwise set $t := t + 1$. Return to Step 2.

For the cluster optimisation problem, the TDE algorithm remains the same except that each point in S_1 or S_2 represents a cluster of atoms with corresponding energy or function value. The crossover operator (13) is adjusted accordingly. In particular, the j component of the vector y^i is now the j atom of the cluster y^i.

3.3   A genetic algorithm for structure calculations

General description
A GA is a global optimisation procedure that uses an analogy of the genetic evolution of biological organisms [21, 22]. It is an heuristic search procedure that modifies function values encoded as binary strings, through the application of predefined reproduction operators in a stochastic manner. The binary string, referred to as a chromosome, is divided into individual sections called genes. The algorithm has some similarities with the TDE method and pseudocode for the genetic algorithm is shown below.

Pseudocode for the genetic algorithm

Step 1 Initialise population() of size $npopsize$

Step 2 while[not converged] do

Step 3 assign population fitness()

Step 4 do from 1 to $npopsize$

    select parents(P_1,P_2)

    reproduction(P_1,P_2,child)

    enddo

Step 5 select next generation()

Step 6 endwhile

There is an absolute measure of quality also assigned to an individual called the fitness f that is a function of the genes. When designing a GA to search for the molecular structure with minimum energy an individual x_a from the population corresponds to a molecular configuration and a gene is the atomic co-ordinates of an atom, binary encoded. The chromosome of x_a is the string of genes defining the molecular structure and the fitness measure of x_a is the potential energy.

Following the initial random generation of a population, the reproductive phase of the GA starts by selecting two parents P_1 and P_2. This choice is weighted depending on the fitness of the parents. With two parents selected the binary digits which encode the string are changed in part by crossover and mutation (i.e. genes are interchanged). Crossover in its simplest form replaces some of the digits that encode parent P_1 by those of parent P_2. Mutation also changes some of the replaced digits with a small probability and ensures that the likelihood of exploring any subset of the search space is always non-zero. The next population is selected when there is an equal number of 'adult' and 'child' individuals.

Parent selection

In a population P of N parents, $P = \{x_1, ...x_N\}$, let the individuals be ordered in descending fitness so that $f(x_i) < f(x_{i+1})$ for $i = 1, ..., N - 1$. Further, let p_i denote the probability that a parent x_i is selected as one of the two parents to which the reproductive operators are applied. The probabilities $\{p_1, ...p_N\}$ such that $\sum^N p_i = 1$, denotes the set of parent selection probabilities which are assigned by a parent selection scheme. Our procedure is to use a binary-tournament selection, first introduced by De Jong [21]. This method takes a random choice of 2 individuals from the population P and selects the fitter as the parent. The selection probabilities $p_i, \quad i = 1, ..., N$ are given by $p_i = (2(N - i) + 1)/N^2$, for $i = 1, ..., N$ where $\sum^N p_i = 1$. This scheme assigns a selection probability to an individual according to its relative fitness, so that the most fit individual has a probability $p_1 = (2(N - 1) + 1)/N^2$, the second most fit $p_2 = (2(N - 3) + 1)/N^2$ and so on down to $p_N = 1/N^2$ but without requiring that the individuals be sorted by fitness.

Reproduction

To apply the GA in the optimisation of single element molecular structures it is necessary to give a definite meaning to each of the terms defined in the general description. The population refers to the set of isolated N atom molecules. The fitness of an individual is the molecule's total potential energy and the genes define the molecular structure. It is also necessary to assign

a potential energy per atom and the genetic encoding procedure should be such that only those atoms j which most influence the potential energy of atom i are located in proximity on the chromosome. The reproduction process continues in four separate stages (a) parameter encoding, (b) crossover, (c) mutation and (d) structural relaxation. Another advantage of the GA is that the method will easily parallelise. The child production and structural relaxation from two parents can in principle be carried out on single processors in parallel. This is the most time-consuming part of the algorithm, the choice of parents from a given population being a very fast process. However in this paper all the calculations were carried out on serial computers.

Basis of the algorithm

The optimisation scheme replaces a high energy region, normally less than half of the molecule, (referred to as h) of parent P_1 with a low energy region (referred to as l) of parent P_2. The algorithm also works by interchanging regions that are random and not selected on the basis of high or low energy. In this case convergence is slower but in some rare cases better optima are found when the energy method locates only a local minimum. For the energy selection process both regions from the parent molecules must contain the same number of atoms. The high (or low) energy regions were selected as follows. The use of empirical potentials allows the definition of a potential energy per atom. Thus a high (or low) energy atom in the whole cluster is first selected. The 2-tournament selection process was used to determine this atom. This involved randomly selecting a subset of the cluster atoms pairwise, comparing energies and rejecting the atom in the pair with the lowest (highest) energy until only one atom remained. Next a plane is chosen with randomly selected position and orientation but such that h region of parent P_1 is on one side of the plane and the centre of mass $\mathbf{c}$ of P_1 lies on the other side. The second molecule P_2 is now rotated so that it also is similarly orientated with l on one side of the plane and its centre of mass on the other. The genes corresponding to the h and l regions are then interchanged. The practical implementation of this step entails first choosing a location on a molecule biased towards a high or low energy region. The vector $(\mathbf{h} - \mathbf{c})/|\mathbf{h} - \mathbf{c}|$, where $\mathbf{h}$ is usually chosen as the atomic co-ordinates of the highest energy atom in the region h, (or sometimes random co-ordinates within a bond length of the high energy atom) forms the unit normal for the dividing plane. This dividing plane is then located between $\mathbf{c}$ and $\mathbf{l}$, where $\mathbf{l}$ is a vector located in region l, on the other molecule in such a way as to preserve the number of atoms. This number is not fixed but varies between 1 and $N/2$ for N-atom molecules. A typical value of N used for the results given in Table 1 was $N = 30$. The dividing plane being located at $\mathbf{c} + r(\mathbf{l} - \mathbf{c})$ where r is a uniform random variate. Parent P_2 is rotated to align ($\mathbf{h}$ - $\mathbf{c}$) with ($\mathbf{h}$ -l) and the region h is replaced by l by a simple translation. For a two component system, the procedure is a little more complicated as the number of atoms of both species has to be preserved. Consider the first parent P_1 and let there be n_1 atoms of species 1 and n_2 atoms of species 2 in region h. The low energy region l of parent P_2 is chosen as before to contain $n_1 + n_2$ atoms. However there is no guarantee that n_1 and n_2 are the same in both h and l. Thus the random choice of l in parent P_2 is repeated up to a maximum of n_{max} times. If a configuration does not arise with n_1 and n_2 the same in both h and l in these n_{max} attempts then a strategy is implemented to change the atoms types in l. It was found that when applying the method to $C - H$ systems with $n_{max} = 10$ an atomic arrangement could be found on 70% of occasions where n_1 and n_2 were the same in both h and l. In the 30% of occasions where this was not the case, the book-keeping strategy adopted was randomly to replace additional H atoms, preferentially those with the highest co-ordination, with a C atom but in the case of additional C atoms to replace with H atoms preferentially those of lowest co-ordination.

Single point crossover

Crossover assembles a new child geometry from the low energy parts of both regions. This
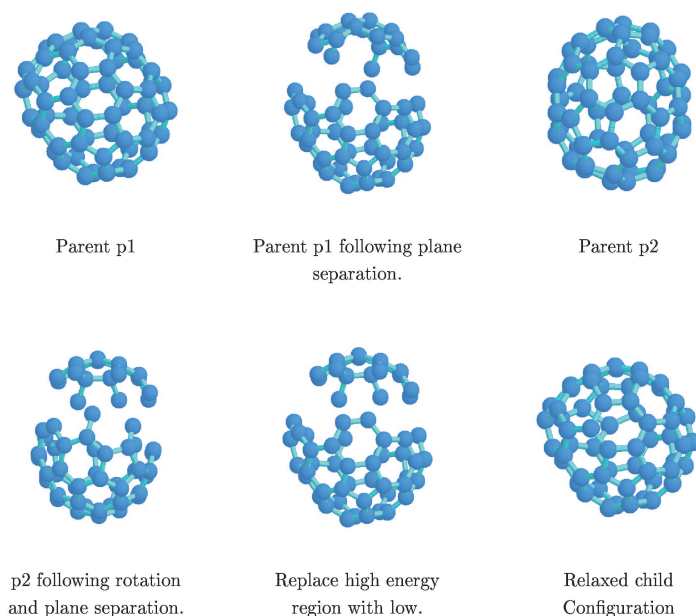
Figure 2: The GA implementation for molecules.

process is shown in figure 2. The crossover takes place at the point separating the genes of d_1 and d_2. The child chromosome takes the d_1 genes of P_2 with the d_2 genes of P_1. In terms of atom co-ordinates, the child molecule takes the geometry of its first parent except for a high energy region which is replaced with a low energy region from the other parent. The number of atoms replaced by this operation varies from 1 to n_2 depending on the value of r.

Mutation

Mutation occurs by replacing typically one or two atoms of high energy (poor placement) whilst maintaining the total number of atoms in the molecule. The replaced atoms are not already in the high energy region and are removed completely from the low energy region of parent P_1. This means that the number of atoms added in the low energy region during the crossover must be increased proportionately. We choose an effective mutation operator that randomly replaces the genes that encode the placements of up to k atoms of the child chromosome with those from the parent P_2. The procedure is as follows. On the new child chromosome, locate $i = 1$ or 2 or ....k atoms with high energies $h_1, h_2, ...h_k$. Mutation occurs when $i > 0$ and then the genes corresponding to the atoms $1, ..., k$ are deleted. New genes are added by effectively translating the second parent normal to the separation plane until k more atoms lie within the appropriate region. Typically $k = 0, 1$ or 2. For the two component $C - H$ system, the strategy is a little different. Locate only $i = 1$ or 2 or ....l, H atoms with poor placement. Locate a similar set of

C atoms that have a low number of neighbours and reassign the l H atoms randomly to points in the neighbourhood of these C atoms.

Child structure relaxation

The child structures are relaxed using a local optimisation routine BFGS developed by Liu and Nocedal [23]. This has the advantage that the memory requirements and iteration time scale linearly with the dimension of the search space and in addition any second derivative discontinuities in the potential energy surface do not present problems. A fixed tolerance is set and this algorithm is iterated until the sum of squared forces is less than this tolerance.

Population replacement strategy

The GA pseudocode suggests that every generation creates N offspring from the current population of N adults. The question that arises is how to select the replacement generation from this set of $2N$ parent and child structures. The strategy adopted here is to select the subset of the N most fit individuals from the current set of adults and offspring. This strategy ensures that the increase in $f(x_1)$ is monotonic, where x_1 is the most fit individual in the population. To counteract premature convergence a restarting procedure was employed. If the chromosomes of the population become identical to within a small Hamming distance (typically 5%) in insufficient iterations to expect global convergence to have occurred, then the population is reinitialised but keeping the chromosome of the fittest individual.

The main difference between the genetic algorithm applied to the optimisation of clusters and the TDE method is that the TDE method contains no physical information about the system. It samples function space rather than the space in which the atoms are distributed. On the other hand the genetic algorithm uses a physical principle of cut and paste based on the physical space. As a result it is much more effective at locating the global minima. If the genetic algorithm were based on the function space instead then it would perform poorly in comparison with the TDE method. Thus an important principle in constructing algorithms for global optimisation problems is to use all the available information about the system rather than to treat the problem as one of optimising a function about which nothing is known.

4   Implementation and comparisons of the algorithms

The search region Ω and initial atom positions

The search region for all the problem is constructed in the following way in a Cartesian system (x, y, z). Distances are measured in Å. The first atom is fixed at the origin. The second atom has $y = z = 0$. The x co-ordinate position due to the second atom is taken as $x_1 \in [0, 4]$. The third atom has $z = 0$ and the second and third variables are such that its x and y co-ordinates are given by $x_2 \in [0, 4]$ and $x_3 \in [0, 4]$. The coordinates of any other atom i are taken randomly to lie in

$$\left[-4 - \frac{1}{4} \left\lfloor \frac{i-4}{3} \right\rfloor, \ 4 + \frac{1}{4} \left\lfloor \frac{i-4}{3} \right\rfloor \right],$$

where $\lfloor r \rfloor$ is the nearest least integer with respect to $r \in I\!R$. For instance, three variables (x_4, x_5 and x_6) due to the 4-th atoms lie in $[-4, 4]$ and the variables for the 5-th atoms in $[-4.25, 4.25]$ and so on. The entire region Ω for N atoms will therefore be

$$\left[-4 - \frac{1}{4} \left\lfloor \frac{N-4}{3} \right\rfloor, \ 4 + \frac{1}{4} \left\lfloor \frac{N-4}{3} \right\rfloor \right].$$

When the GA was applied to the large clusters a random initial distribution of atoms in Ω was chosen.

12

The initial population size: It was found necessary to use a larger population size for the TDE compared to the GA. For the TDE a population size of $N = 5n$ was found to be necessary, where n is the number of dimensions of the problem ($n = (3N_a - 6)$ where N_a is the number of atoms). For MSL and TMSL $5n$ random points were also generated at the begining of each iteration. However good convergence could be achieved even for large clusters using the GA with a population size of only 30. However for comparison purposes with the small clusters the same initial population size $N = 5n$ was also used for GA. The results generated with the larger clusters were obtained with the more optimal value of $N = 30$, which reduces the number of function evaluations considerably.

Other parameters: TMSL and TDE require both the number of points required to calculate the topographical cluster and the number of nearest neighbours k_g used to determine the centre of the cluster. A value of $k_g = 4$ was used throughout. A fraction of points e.g. $N_g = \gamma_1 j N$ used in TMSL, where j is the iteration counter and γ_1 was normally taken as 0.15. TDE maintains a fixed N throughout all iterations and a value of $N_g = 0.2N$ was used. The value of t_{max} used in the TDE method was set to 4 (see Step 2, Page 8) and $M = 2n$. The same stopping condition was used for all algorithms in Table 1. A value of F of 0.5 was found to be optimal. Other information concerning the parameters of MSL and TMSL can also be found in [24].

A comparison of algorithms using L-J clusters: We compare TDE and GA with the multi-level single linkage (MSL) and topographical multi-level single linkage (TMSL) algorithms mentioned in the introduction. In a earlier study [24] it was found that MSL and TMSL are the best performers within a number of global optimisation algorithms, e.g. controlled random search and simulated annealing. We first compare the algorithms in Table 1 using the average number of function evaluations (fe) and the best minimum value obtained by the algorithms. Average fe has been obtained over 10 independent runs. The result in bracket is the number of times the the global minimum was obtained. N_a in Table 1 represents the number of atoms. We have used a Pentium 4 machine with 256 MB ram and 1600 MHz Processor.

Table 1: Comparison using the Lowest Minima found

N_a	MSL Energy (eV)	fe	TMSL Energy (eV)	fe	TDE Energy (eV)	fe	GA Energy (eV)	fe
3	-3.00(10)	1534	-3.00(10)	1326	-3.00(10)	3314	-3.00(10)	1550
4	-6.00(10)	3479	-6.00(10)	1874	-6.00(10)	4772	-6.00(10)	3673
5	-9.10(6)	12479	-9.10(10)	10721	-9.10(10)	11115	-9.10(10)	8032
6	-11.32(0)	59209	-12.30(1)	45362	-12.30(5)	59340	-12.30(10)	31395
7	-16.50(5)	131743	-16.50(4)	143733	-16.50(10)	480450	-16.50(10)	48900
8	-16.48(0)	173960	-18.05(0)	186836	-19.81(10)	773190	-19.82(10)	121247
9	-19.41(0)	595712	-19.05(0)	556624	-24.10(10)	3315810	-24.11(10)	346397
10	-19.75(0)	2783247	-19.90(0)	842733	-28.42(10)	7346875	-28.42(10)	721370

Results presented in Table 1 show that in terms of fe TMSL and GA is comparable. However, as the number of atoms increases robustness of TMSL in terms of locating the best minimum value falls off. GA was able to locate the best known minimum values for all cases. TDE was able to find the best known minimum for all cases except $N_a = 6$. Moreover, TDE is the worst performer in terms of fe. In Table 2, we compare the cpu times of the algorithms corresponding to the results in Table 1.

N_a	MSL	TMSL	TDE	GA
		Table 2: Comparison using cpu time		
3	0.12	0.14	1.27	0.04
4	0.77	0.76	0.64	0.47
5	0.63	1.68	1.33	1.27
6	3.96	6.93	2.03	1.93
7	9.32	19.33	6.21	3.89
8	22.50	42.43	41.93	8.60
9	71.17	94.99	75.60	23.63
10	83.70	134.72	181.78	60.22

It is clear from Table 2 that GA is the best algorithm in terms of cpu time, even without using the optimised population size. The worse performers are the TMSL and TDE algorithms. These two algorithms need high cpu times due to the calculation of 'graph minima' at each iteration.

Argon and silver clusters: For the problem of finding the minimum energy configurations, the GA always outperformed the TDE method. In fact other global optimisation methods such as controlled random search and simulated annealing were also tested and these were also found to be inferior to the GA. An example, comparing the GA and TDE methods is also given in the section on silicon clusters below. However there are some interesting geometrical structures that arise with close-packed configurations. Both the embedded atom potential and the Lennard-Jones potential result in close-packed structures and so we consider the geometries of these together in this section. These geometries show some interesting features as the number of atoms in the system changes. In this section all the results were calculated using the GA. Figure 3 shows some of these geometrical configurations for clusters containing 7, 38, 55, 147 and 400 atoms. The 7, 38 and 147 atom clusters have icoshedral symmetry whereas the 55 atom cluster is a structure that is often used in nanotechnology, usually with gold atoms for a number of interesting technological applications. The 400 atom cluster shows that when the number of atoms is large close-packed facets along the edges of the clusters can be seen as the cluster begins to take on the appearance of the bulk material.

The convergence of the GA is only partially dependent on the symmetry of the cluster. For example one might expect that the symmetric clusters shown in figure 3 would be determined quickly by the GA compared to non-symmetric clusters. This is in fact only partially the case. Numerical experiments carried out for clusters with atoms ranging from 35-40 show that the number of generations required to determine the minimum energy structure for the symmetric cluster is not out of line with its neighbours. We have carried out five independent runs. The difference is only the number of generations required from different starting conditions, not the value of the minimum. These are shown in Table 3, where AGV represents the average number of generations. Although with any optimisation algorithm, one can never be certain that the global optimum has been reached, the minimum energy values given in table 3 are in line with what would be expected. It is worth noting that because of the large number of nearby local minima that the algorithm takes a large number of generation to fine tune. For example for the 40 atom cluster a value of -100.27 eV was obtained after 486 generations of the algorithm in a particular run. It took a further 641 generations to reduce the energy by 0.02 eV to the final value.

Silicon clusters: The GA and TDE can be successfully applied to the Tersoff silicon potential for small clusters. Figure 4 shows the geometry of small Si clusters calculated using parameterisation (A). In this case cages are predicted for upwards of 4 atom clusters. The 6 and 7 atom structures predicted are in agreement with recent experimental results [25, 26]. An interesting

Table 3: Minimum energies of Ag clusters, calculated using an embedded atom potential.

No. of atoms in cluster	35	36	37	38	39	40
AVG	526	468	827	469	732	1118
Cluster potential energy (-eV)	86.70	89.55	92.34	95.40	97.78	100.29

example concerning the evolution of the Si_{18} atom cluster is that two isolated Si_8 and Si_{10} atom structures are predicted to have a lower energy with this parameterisation. The evolution of this cluster using the GA is shown in figure 5. The minimum energy of -81.46 eV was determined after approximately 500 generations of the algorithm starting from an initial population size of 32. The dissociation of an 18 atom cluster occurs because the silicon potential is many-body and depends on the local co-ordination and bond angles and not just on the proximity of the atoms. Such a dissociation would not occur with a Lennard-Jones or EAM potential. Although the global minimum energy configurations for the Si_{18} cluster with the Tersoff potential is two separated clusters, Goedecker and et.al. [27] have found a Si_{18} cluster using density functional theory and a new 'dual minima hopping method' to locate minima, which is a single entity and not divided. They did not calculate the energies of the isolated Si_8 and Si_{10} clusters to see if these produced a lower energy. In any event the Tersoff potential is known to be less accurate than density functional theory for small clusters as it was fitted to bulk material properties. There is no reason other than available computing power why the genetic algorithm might not be applied to ab initio calculations such as those described in [27].

7 atoms   38 atoms

55 atoms   147 atoms
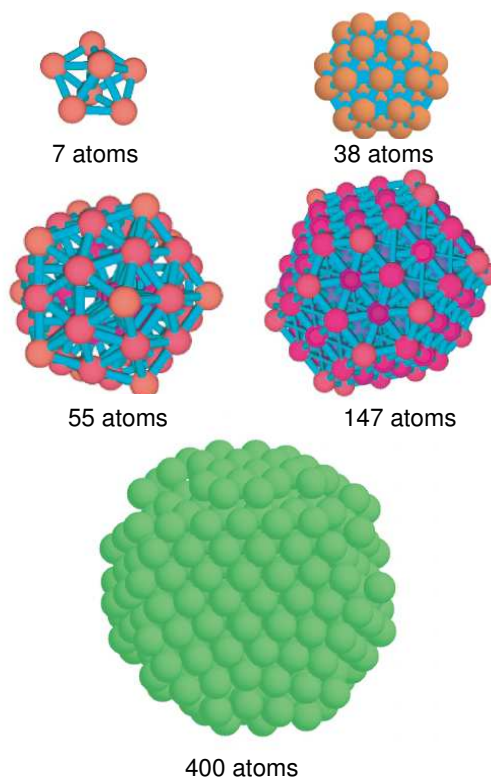
400 atoms

Figure 3: The geometry of the closed-packed structures containing (a) 7, (b) 38, (c) 55, (d) 147 and (e) 400 atoms.

Larger clusters containing between 50 and 58 atoms are shown in figure 6, calculated using parameteristaion (B). Here there are internal sp^3 bonds unlike the carbon potential. The 60 atom Si structure is about 3.3 eV more energetically favourable than the corresponding fullerene cage [28]. These structures are very similar to recent calculations using ab-initio potentials [29].

Carbon clusters

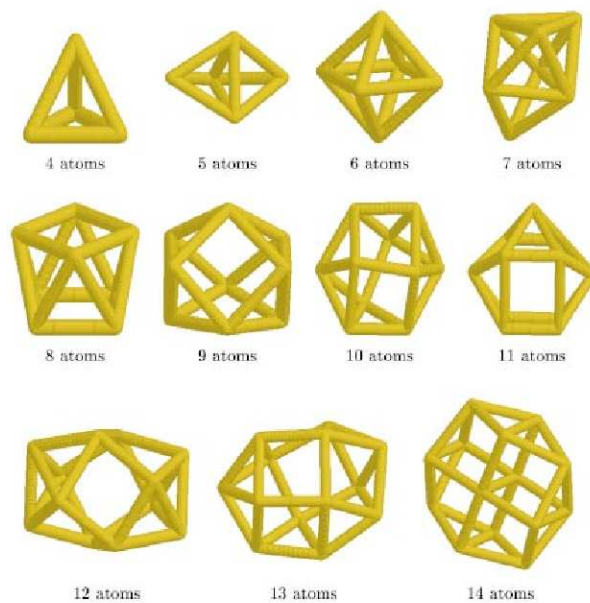Table 4 gives a list of the optimum carbon structures calculated using the GA for the Brenner carbon potential.

Figure 4: Small S_i clusters calculated using parameterisation (A).

Number of Atoms	Potential Energy of Molecule (eV)	Energy / Atom (eV)	Geometry
3	-12.40	-4.1329	l
4	-18.58	-4.6439	l
5	-26.33	-5.2651	m
6	-33.96	-5.6600	m
7	-41.10	-5.8714	m
8	-47.91	-5.9883	m
9	-54.50	-6.0551	m
10	-60.95	-6.0950	m
11	-67.32	-6.1199	m
12	-73.61	-6.1342	m
13	-79.90	-6.1416	m
14	-86.11	-6.1507	m
15	-92.34	-6.1560	m
16	-98.61	-6.1631	m
17	-104.80	-6.1647	m
18	-112.29	-6.2386	c
19	-118.83	-6.2543	c
20	-128.39	-6.4195	f
21	-134.52	-6.4058	c
22	-142.09	-6.4586	c
23	-148.41	-6.4528	c
24	-157.16	-6.5483	f
25	-163.13	-6.5252	c
26	-171.98	-6.6146	f
27	-178.01	-6.5930	c
28	-186.88	-6.6743	f
29	-192.67	-6.6438	c
30	-200.44	-6.6813	f
31	-207.00	-6.6774	c

Table 4 : Potential Energy of Lowest Minima (Brenner)

Number of Atoms	Potential Energy of Lowest Minima		
	Potential Energy of Molecule (eV)	Energy / Atom (eV)	Geom-etry
32	-216.66	-6.7706	f
33	-221.78	-6.7206	f
34	-230.30	-6.7735	f
35	-236.38	-6.7537	f
36	-245.45	-6.8181	f
37	-251.23	-6.7900	f
38	-259.97	-6.8413	f
39	-265.69	-6.8126	f
40	-274.64	-6.8660	f
41	-280.38	-6.8385	f
42	-289.34	-6.8890	f
43	-295.12	-6.8633	f
44	-304.09	-6.9111	f
45	-309.02	-6.8671	f
46	-318.54	-6.9248	f
47	-324.06	-6.8949	f
48	-333.38	-6.9454	f
49	-339.03	-6.9190	f
50	-348.37	-6.9674	f
51	-353.88	-6.9388	f
52	-362.83	-6.9742	f
53	-367.92	-6.9419	f
54	-377.69	-6.9943	f
55	-383.08	-6.9789	f
56	-392.37	-7.0066	f
57	-397.73	-6.9777	f
58	-407.25	-7.0216	f
59	-412.74	-6.9956	f
60	-422.55	-7.0425	f

l–linear, m–2D ring, c–cage, f–fullerene

Figure 7 shows also the geometries of some of the small cluster structures including fullerenes. The transition from linear chains to rings occurs between 4 and 5 atoms and the transition from rings to cages between 17 and 18 atoms. In all cases the Brenner potential gave fullerene or fullerene-like cages as the minimum energy configurations for all molecules containing between 18 and 60 atoms. More details of these structures are given in [30].

Hydrocarbon clusters

The modified GA for two component systems is applied to determine low energy hydrocarbon structures. The Brenner hydrocarbon potential has been fitted to a number of small alkane, alkene, alkyne, radical and aromatic molecules and additionally has been shown to reproduce the correct geometry and binding energies of larger hydrocarbon isomers. In this section we test whether the potential identifies the ground state configurations of $C_n H_m$ for specific values of m and n. The optimised structures are compared with molecular data available from the pdb databank at Okanagan University [31]. The GA identifies ground-state isomers with energies in
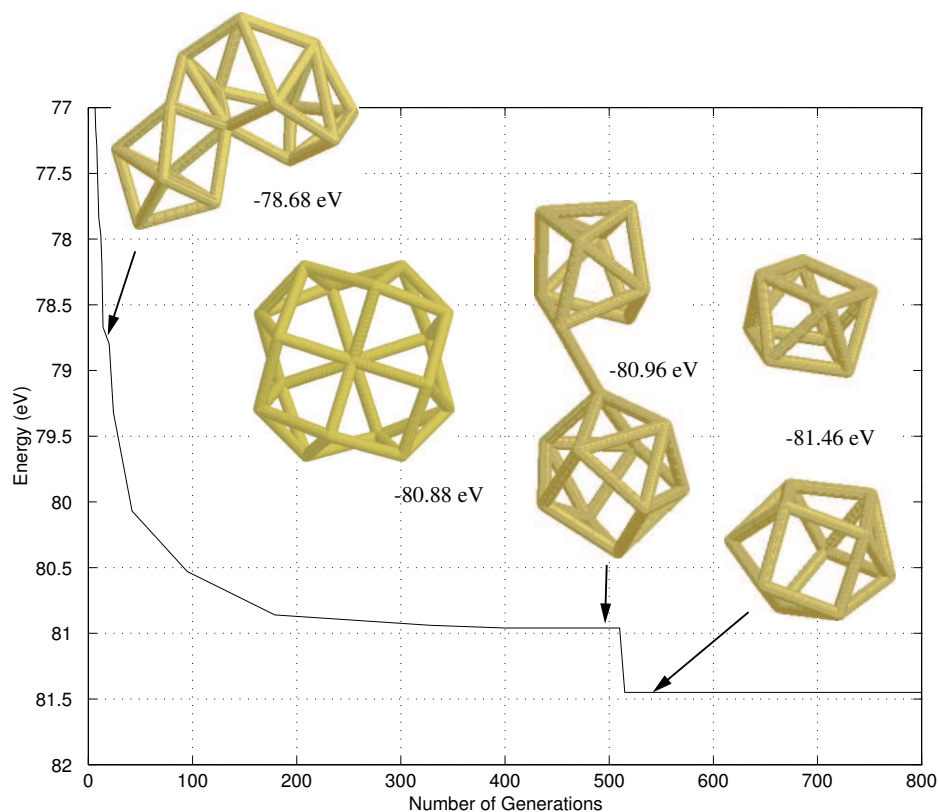
Figure 5: Evolution stages of $18Si$ atoms. After 500 generations the cluster dissociates into Si_8 and Si_{10} geometries. Two Si, clusters have a higher energy at -81.316 eV. The TDE method gives a cluster with a higher energy of -78.25 eV.

every case that are less than or equal to those of the known isomers from the pdb database. The results for a number of alkanes, alkenes, alkynes and aromatic molecules are shown in figures 8-11. Figure 8 shows the application of the algorithm to the Brenner hydrocarbon potential for the alkane hydrocarbon isomers containing relatively small numbers of atoms, figure 9 gives the alkenes, figure 10 the alkanes and figure 11 the aromatics. Data for clusters containing a larger number of atoms is given in table 5. When N_a is greater than 5, the GA determined global geometries that seldom coincide with any contained in the pdb database. These GA determined geometries however produced energies that were lower than those from the molecules reported in the pdb data base. This shows that the Brenner potential cannot be used with confidence to model such structures with $N_a > 5$. The data shows that the stereoisomers are equally energetic with the molecules in the database see for example, C_4H_8. For the aromatic isomers, the GA finds the minimum energy geometries for benzene, toluene, dimethylbenzene and napthalene, although for the case of C_8H_{10}, a configuration with adjacent as opposed to opposite methyl units has also the same energy. However C_8H_6 and C_8H_8 have different molecular structures, with fused pentagonal rings and an eight-membered ring exhibiting preferential stability to ethenylbenzene and ethynylbenzene.
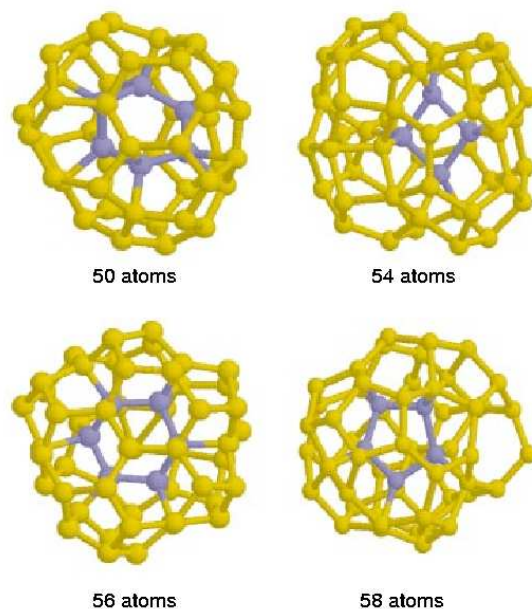
Figure 6: Four examples of Si cage formations with an internally bonded four-fold co-ordinated core calculated using parameterisation (B). Within the core it is possible to see atoms arranged in the approximate forms of a square Si_{54} pentagon Si_{58} or hexagon Si_{50}, Si_{56}. For the cages containing between 48 and 60 atoms, two-thirds of the atoms are three-fold co-ordinated with the majority of the rest having four-fold co-ordination. For 60 Si atoms the cage with the internal core is 3.3 eV more stable than the fullerene. Four examples of Si cage formations with an internally bonded four-fold co-ordinated core calculated using parameterisation (B). Within the core it is possible to see atoms arranged in the approximate forms of a square Si_{54} pentagon Si_{58} or hexagon Si_{50}, Si_{56}. For the cages containing between 48 and 60 atoms, two-thirds of the atoms are three-fold co-ordinated with the majority of the rest having four-fold co-ordination. For 60 Si atoms the cage with the internal core is 3.3 eV more stable than the fullerene.
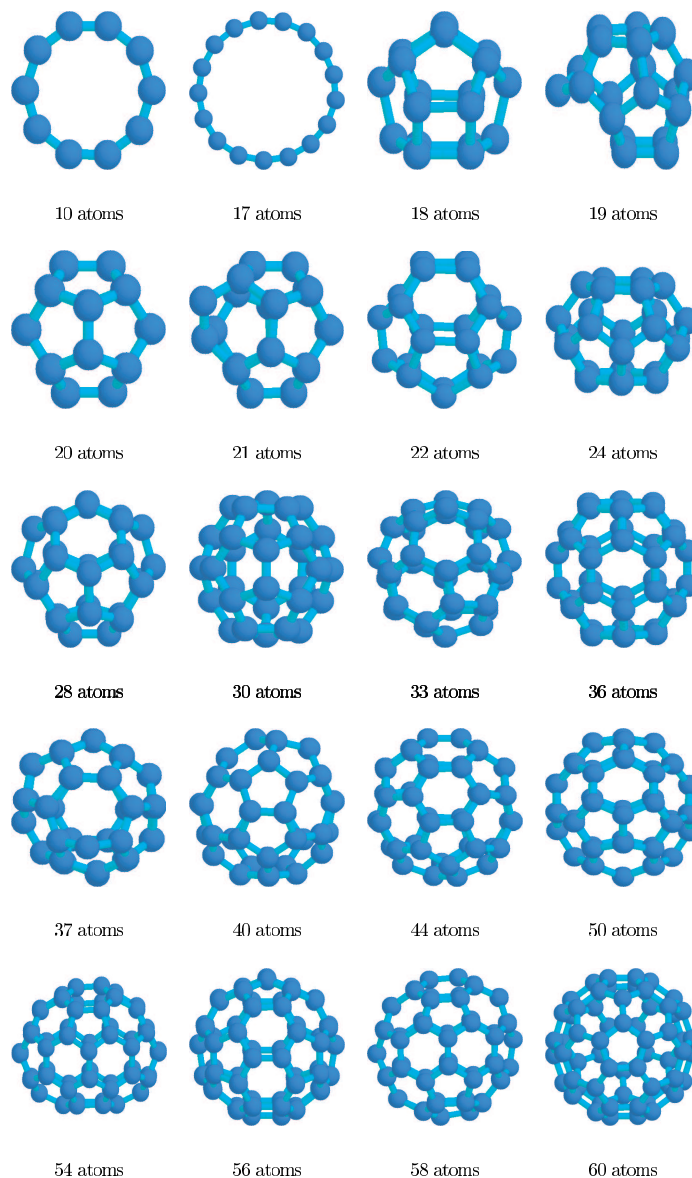
Figure 7: Small carbon molecular structures containing between 10 and 60 atoms calculated using the GA and the Brenner potential.

Table 5 compares the energy of database molecules with those evolved by the genetic algorithm. All configurations are locally optimised with the LBFGS routine according to the Brenner hydrocarbon interatomic potential energy function. The Database ID column provides the filename for the pdb format data base molecule. Note that although the high level of significant figures in the energy table is physically meaningless, is necessary as two molecules with completely different topologies may have energies that are similar to within as little as 0.005 eV.

The fact that the GA can determine lower energy structures than those observed experimentally illustrates the difficulty of fitting real chemical systems with empirical potentials. Whereas this might highlight the inadequacies of the Brenner potential it raises the intriguing possibility that new molecules could be discovered computationally by the application of the GA with accurate interatomic potential energy functions for multi-component systems.
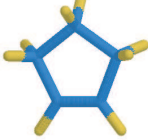
Row 1:	Geometries and Energies for known Alkane Hydrocarbons Isomers
Row 2:	GA optimised geometries for Alkane Hydrocarbons

C_4H_{10}	C_5H_{12}	C_6H_{14}	C_7H_{16}	C_8H_{18}
I-butane	Neopentane	Hexane	Heptane	Iso-octane
-54.27 eV	-66.79 eV	-78.81 eV	-91.08 eV	-103.61 eV
-54.27 eV	-66.79 eV	-79.06 eV	-91.33 eV	-103.86 eV
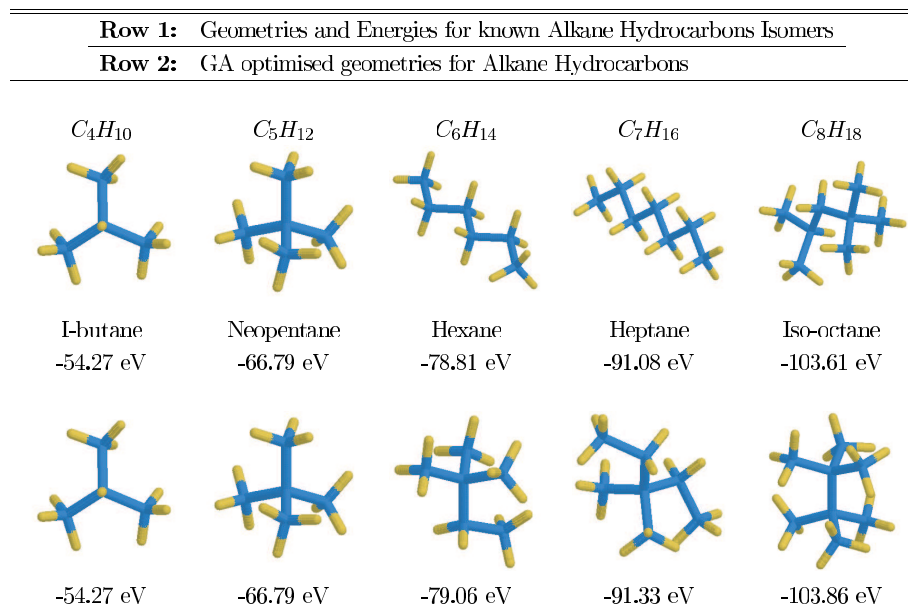
Figure 8: Alkane hydrocarbon structures calculated using the GA and the Brenner hydrocarbon potential.

Row 1:	Geometries and Energies for known Alkene Hydrocarbons Isomers
Row 2:	GA optimised geometries for Alkene Hydrocarbons

C_4H_8	C_5H_{10}	C_6H_{12}	C_7H_{14}	C_8H_{16}
cisbutene	cyclopropane	2,3-Dimethylbut -2-ene	cis 1,3-dimethyl cyclopentane	1,1-dimethyl cyclohexane
-48.92 eV	-61.35 eV	-78.78 eV	-85.90 eV	-98.42 eV
-48.92 eV	-61.35 eV	-78.78 eV	-86.05 eV	-98.42 eV
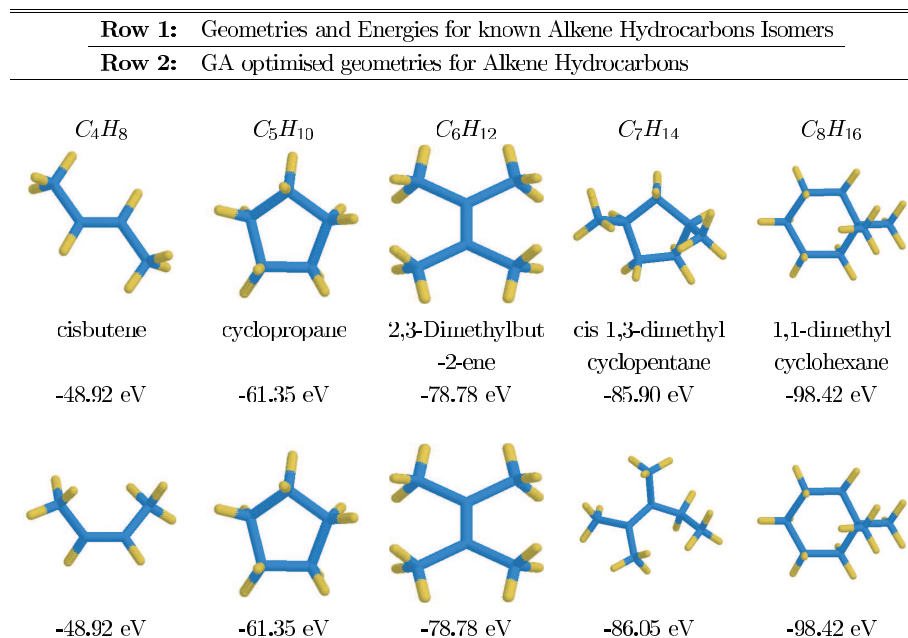
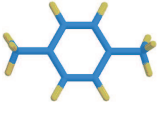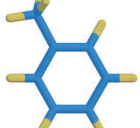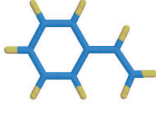Figure 9: Alkene hydrocarbon structures calculated using the GA and the Brenner hydrocarbon potential.
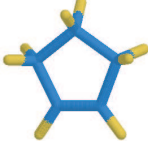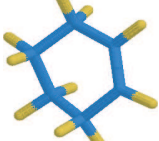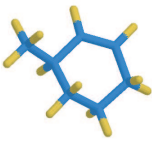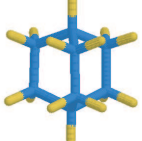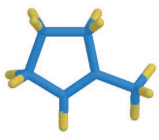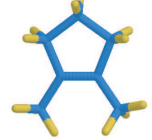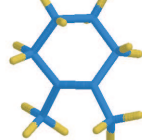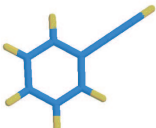
Row 1:	Geometries and Energies for known Alkyne Hydrocarbons Isomers
Row 2:	GA optimised geometries for Alkyne Hydrocarbons

C_4H_6	C_5H_8	C_6H_{10}	C_7H_{12}	C_8H_{14}
ethane	cyclopentene	cyclohexene	3-methylcyclo hexane	Bicyclo octane
-42.24 eV	-55.83 eV	-68.28 eV	-80.55 eV	-93.00 eV
-42.24 ev	-55.83 eV	-68.29 eV	-80.75 eV	-93.14 eV

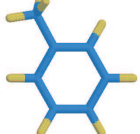Figure 10: Alkyne hydrocarbon structures calculated using the GA and the Brenner hydrocarbon potential.
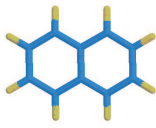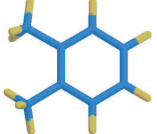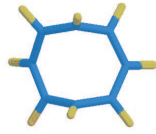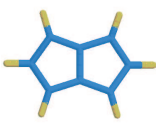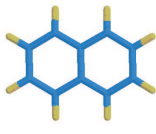
Row 1:	Geometries and Energies for known Aromatic Hydrocarbons Isomers
Row 2:	GA optimised geometries for Hydrocarbons with known aromatic isomers

C_8H_{10}	C_7H_8	C_8H_8	C_8H_6	$C_{10}H_8$
dimethylbenzene	Toluene	Ethenylbenzene	Ethynylbenzene	Naphthalene
-82.33 eV	-69.90 eV	-76.35 eV	-68.32 eV	-91.39 eV
-82.33 eV	-69.90 eV	-76.71 eV	-70.71 eV	-91.39 eV

Figure 11: Aromatic hydrocarbon structures calculated using the GA and the Brenner hydrocarbon potential.

Series	Hydro-Carbon Molecule	No. of Entries in Database	Database pdb ID	Energy, eV		Identical Geome tries
				Database Minimum	GA Minimum	
	C_9H_{20}	3	nonane	-115.620	-116.127	×
	$C_{10}H_{22}$	3	decane	-127.891	-128.144	×
A	$C_{11}H_{24}$	3	34e-3m-6	-140.419	-140.923	×
L	$C_{12}H_{26}$	1	c-7ane	-152.448	-152.935	×
K	$C_{13}H_{28}$	2	b-8ane	-164.719	-165.462	×
A	$C_{14}H_{30}$	2	224477m8	-177.736	-177.742	×
N	$C_{15}H_{32}$	3	b-7ane	-190.013	-190.254	×
E	$C_{16}H_{34}$	3	c-9ane	-201.532	-202.446	×
	$C_{17}H_{36}$	1	d-10ane	-214.048	-214.802	×
	$C_{18}H_{38}$	1	c-10ane	-226.073	-227.164	×
A	C_9H_{18}	6	f-cy5ane	-110.689	-110.689	√
L	$C_{10}H_{20}$	5	bucyclo6	-122.968	-123.219	×
K	$C_{11}H_{22}$	3	g-cy6ane	-135.241	-135.482	×
E	$C_{12}H_{24}$	1	al0557	-147.406	-147.760	×
N	$C_{13}H_{26}$	3	b-6ane	-159.530	-159.935	×
E	$C_{24}H_{48}$	1	al3063	-294.422	-294.633	×
	C_9H_{16}	14	spiro-44	-105.504	-105.544	×
A	$C_{10}H_{18}$	6	s45ane	-117.779	-117.780	×
L	$C_{11}H_{20}$	5	s55ane	-130.055	-130.297	×
K	$C_{12}H_{22}$	4	a-s45ane	-142.328	-142.761	×
Y	$C_{13}H_{24}$	3	b-s45ane	-154.598	-154.840	×
N	$C_{14}H_{26}$	1	al0066	-167.130	-167.152	×
E	$C_{16}H_{30}$	1	a-s56ane	-191.417	-192.094	×
	$C_{17}H_{32}$	2	al0524	-202.017	-203.821	×

Table 5 : Lowest Energy Hydro-Carbon Chart

5 Conclusions

Different algorithms have been compared for the important problem of determining the most stable configuration of a system of particles in space which interact through classical force-fields which are both attractive at large separation and repulsive when the particles lie close together. Modern global optimisation algorithms have been shown to obtain some of the correct configurations when the number of particles is small but the GA outperforms these for larger systems.

The algorithm would also in principle be suitable for application with ab initio methods although the determination of the total system energy would be more time-consuming. The algorithm can also be easily parallelised since the time consuming part of the calculation is that involved with the local optimisation of each cluster in the population.

The algorithm is also robust in that it can determine the minimum energy configuration of a system of particles without any pre-knowledge of what the structure should be. All that is required is a reasonably sized box in which the particles are distributed. In contrast, many chemists use intuition followed by local optimisation. This can often fail as demonstrated in [13] where the genetic algorithm found more optimal minima than had previously been determined by enlightened guesswork.

The advantage of the GA compared to a black box global optimisation approach is that it is physically based and uses the idea that an already existing cluster of atoms can be rearranged with knowledge of the configurations of other clusters. This evolution strategy, together with the relatively small number of candidate clusters required, gives it a considerable advantage over the MSL and TDE methods.

The GA also found lower energy structures for the hydrocarbons modelled with the Brenner potential than those experimentally observed. The experimentally observed structures always turn out to be local minima in the Brenner description but not always global ones. With empirical potentials it is always difficult to describe structures adequately for which the potential has not been fitted. This illustrates that even the Brenner potential, which is one of the most sophisticated available for modelling carbon and hydrocarbons, cannot describe all cases with accuracy and that ab initio methods should be used wherever possible.

Molecular design and candidate geometries for new drugs are increasingly being investigated using computers. The combination of the GA approach with molecular design techniques might therefore yield substantial benefits in this developing field.

References

[1] Tersoff, J. (1988), New Empirical approach for the Structure and Energy of Covalent Systems, *Physics Review B*, 37:6991-7000.

[2] Tersoff, J. (1988), Empirical Interatomic Potential for Silicon with improved Elastic Properties, *Physics Review B*, 38:9902-9905.

[3] Tersoff, J. (1988), Empirical Interatomic Potential for Carbon with Applications to Amorphous Carbon, *Physical Review Letters*, 21:2879-2882.

[4] Daw, M.S. and Baskes, M. I. (1987), Embedded-atom method: Derivation and applications to impurities, surfaces and other defects in metals *Physical Review B*, 29:6444-6453.

[5] Ackland, G. J., Tichy, G., Vitek, V. and Finnis, M. W. (1987), Simple N-body potentials for the noble metals and nickel *Philosophy Magazine A*, 56:753.

[6] Ashcroft, N.W. and Mermin, N.D. (1976), *Solid State Physics*, Holt, Rheinhart and Winston, pp 398-402.

[7] Carlsson, A. E. (1990), Beyond Pair Potentials in Elemental Transition Metals and Semiconductors, *Solid State Physics*, 43:1-90.

[8] Ali, M.M. and Storey, C. (1997), Aspiration Based Simulated Annealing Algorithm, *Journal of Global Optimization*, 11:181-191.

[9] Rinnooy Kan, A.H.G. and Timmer, G.T. (1987), Stochastic Global Optimization Methods; Part II : Multilevel Methods, *Mathematical Programming*, 39: 57-78.

[10] Ali, M.M. and Storey, C. (1994), Topographical Multilevel Single Linkage, *Journal of Global Optimization*, 5:349-458.

[11] Ali, M.M. and Storey, C. (1994), Modified Controlled Random Search Algorithms, *International Journal of Computer Mathematics*, 54:229-235.

[12] Ali, M. M. and Törn, A. (2000), Optimization of Carbon and Silicon Cluster Geometry for Tersoff Potential using Differential Evolution, in *Optimization in Computational Chemistry and Molecular Biology : Local and Global Approaches*, Edited by A. Floudas and M. Pardalos, Kluwer Academic Publisher, 287-300.

[13] Hobday, S. and Smith, R. (1997), Optimization of Carbon Cluster Geometry Using a Genetic Algorithm, *Journal of Chemical Society, Faraday Transactions*, 93:3919-3926.

[14] Judson, R. (1997), Genetic Algorithms and Their Use in Chemistry, *Reviews in Computational Chemistry*, 10:1-73.

[15] Niesse, J.A. and Mayne, H.R. (1997), Global Optimization of Atomic and Molecular Clusters Using the Space-Fixed Modified Genetic Algorithm Method, *Journal of Computational Chemistry* 18:1233-1244.

[16] Doye, J. P. K., Miller, M. A. and Wales, D. J. (1999), Evolution of the Potential Energy Surface with Size for the Lennard-Jones Clusters, *Journal of Chemical Physics*, 111:8417-8428.

[17] Locatelli, M. and Schoen, F. (2003), Efficient Algorithms for Large Scale Global Optimization: Lennard-Jones Clusters, *Computational Optimization and Applications*, 26:173-190.

[18] Deng, Y. and Rivera, C.(2000), Approximate Energy Minimization for Large Lennard-Jones Clusters, *Journal of Global Optimization*, 16:325-341.

[19] Brenner, D. W. (1990), Empirical Interatomic Potential for Hydrocarbon for Use in Simulating the Chemical Vapor Deposition of Diamond Films, *Physics Review B*, 42:9458-9471.

[20] Brenner, D. W. (1990), Erratum: Empirical Interatomic Potential for Hydrocarbon for Use in Simulating the Chemical Vapor Deposition of Diamond Films, *Physics Review B*, 46:1948.

[21] De Jong, K. A. (1975), An analysis of the behaviour of a class of genetic adaptive systems. *PhS thesis Uinversity of Michigan*

[22] Goldberg, D. E. (1989), *Genetic Algorithms in search, optimization and machine learning*, Addison Wesley.

[23] Lui, D. C. and Nocedal, J. (1989), On the limited memory BFGS method for large scale optimization *Math. Programming*, 45:503-528.

[24] Ali, M. M., Storey, C. and Törn, A. (1997), Application of Stochastic Global Optimization Algorithms to Practical Problems, *Journal of Optimization Theory and Applications*, 95: 545-563.

[25] Raghavachari, K and Rohlfing, C. M. (1988), Bonding and stabilities of small silicon clusters - A theoretical study of Si_7 - Si_{10} *J. Chem. Phys.*, 89:2219-2234.

[26] Raghavachari, K and Rohlfing, C. M. (1992), Structures of Si_{10} - are there conventionally bonded low energy isomers ? *Chem. Phys. Lett.*, 198:521-525.

[27] Goedecker, S, Hellmann, W and Lenosky, T (2005), Global Minimum Determination of the Born-Oppenheimer Surface within Density Function Theory, *Phys. Rev. Lett.*, 95:055501.

[28] Kroto, H. W., Heath, J. R., Óbrien, S. C., Curl, R. F. and Smalley, R. E. (1985), C_{60} : Buckminsterfullerene, *Nature*, 318:162-167.

[29] Rothlisberger, U, Andreoni, W and Parrinello, M. (1994), Structure of Nanoscale silicon clusters, *Phys. Rev. Lett.*, 72:665-668.

[30] Hobday, S. and Smith, R. (1999), Application of Genetic Algorithms and Neural Networks to interatomic potentials, *Nucl. Instrum. and Meth. B*, 153:247-263.

[31] Woodcock, D. (1997), http://www/sci.ouc.bc.ca/chem/programmes/modsrch.htm, Hydrocarbon Protein Data Bank (PDB) database, Okanagan University, Canada.

[32] Locatelli, M. (1998), Relaxing the assumptions of the Multilevel Single Linkage Algorithm, *Journal of Global Optimization*, 13 (1): 23-42.