

Visual Mining and Spatio-Temporal Querying in Molecular Dynamics

Olga Sourina and Nikolay Korolev*

NANYANG TECHNOLOGICAL UNIVERSITY, NANYANG AVENUE,

639798, SINGAPORE

EOSOURINA@NTU.EDU.SG

ABSTRACT

Continuously increasing power of computers brings possibility to model dynamic behavior of the systems studied by experimental nanoscience (single-molecule mechanics, electron and atomic force microscopy) with explicit all-atom presentation by applying various molecular dynamics approximations. Computer simulation in molecular dynamics usually results in gigabytes of time-dependent data. On the other hand, computational and querying tools could help to analyze such information. We propose a novel geometric querying method for data analysis in molecular dynamics. This method is based on a novel geometric model for querying of spatio-temporal data. In our model, data from database or file are interpreted geometrically as multidimensional points with time dimension. A geometric query is a query solid of any shape specified by its parameters, location and time. These queries are formulated with geometric objects and operations over them to form the query solid changing over time. The geometric objects and operations are described with implicit functions. We proposed spatio-temporal queries for data analysis in molecular dynamics that allow formulate new hypotheses. The geometric queries in our model can be considered as motion classifiers. The resulting sets are coordinates of molecules and atoms inside the particular shape changing over time. With the uniform geometric model we integrate visual mining and querying of time-dependent data employing 3D visualization tools.

Key Words: *Molecular Dynamics, Spatio-Temporal Queries, Implicit Functions, Visual Mining*

INTRODUCTION

Computer simulation in molecular dynamics (MD) usually results in gigabytes of time-dependent data that can be stored in spatio-temporal databases or files. Combination of computational MD data with experimental results of nanoscale (single-molecule) mechanical manipulations and microscopic observations provide true “atomic resolution” insight into many phenomena and have intriguing perspective both in understanding important (first of all, biological) processes and in creation of new materials and technologies. To understand and interpret the MD results the user often needs to be interactively involved in the process of data mining and querying of the spatio-temporal data. It becomes more and more important to give the user an easy understanding of the data set. Visualization offers an intuitive way of data analysis that can help to discover data patterns and structures hidden in the data [1]. Data visualization techniques, when incorporated with querying algorithms, could improve interpretability and usability of the time-dependent data and improve analysis of that data. First, the user is interested not only in the results of mining and querying, but also in the data itself, in the querying process, and in spatial relationship inside the data changing over time. Therefore, the user should analyze data visually and be involved into the querying process to make it more efficient and accurate. Thus, visualization techniques could be used not only for visualization of results of querying and data mining but for visual mining and visualization of the querying process. The user should also have an opportunity to select the projection directions for data set. To incorporate visualization techniques, the existing spatio-temporal database systems use the result of querying as the input for visualization system that is costly and inefficient. The better solution is to combine two processes together, which means to use the same model to query and visualize time-dependent data.

Usually we are interested both in the data and in the underlying processes and structures that these data provides. In case of time-dependent data, we often are interested in forecasting future developments based on trends and cyclic behaviors in the data [2]. However, finding such valuable information hidden in data is a complicated task. Visualization offers the user an intuitive way of analysis that can help to discover data patterns and structures. Data visualization techniques when incorporated with data mining algorithms could improve interpretability and usability of the data, and allow the user to come up with new hypothesis. There are different molecular visualization programs. Possibilities for geometry analysis and visualization within existing packages for molecular dynamics such as CHARMM [3, 4], AMBER [5], and GROMACS [6] are rather limited. With VMD [7] and GOpenMol [8] software systems we can display, animate, and analyze large biomolecular systems using 3-D graphics and built-in scripting. But structure geometry analysis still could be improved. In VMD [7], an extensive atom selection syntax is implemented but range queries are limited to sphere region - “find atoms within 6 Å of (1, -2.3, 0)”.

On the other hand, the goal of works in the area of visual mining and querying of spatio-temporal data is to propose data representation model and query model able to handle time-dependent geometries including those changing continuously that describe moving objects [9, 10]. Spatio-temporal predicates are introduced to query time-dependent data [10]. But existing models and implemented software systems do not provide geometric queries of any shapes. We propose a novel geometric querying method for data analysis in molecular dynamics. First, the molecular system is visualized for all time frames. Then, the user analyses the data visually and formulates different geometric hypotheses that can be tested by posing and implementing spatio-temporal queries. In our earlier work [11], we proposed and fully described geometric query model for relational databases with implicit functions [12]. In this paper, we propose the uniform geometric query model to handle spatio-temporal data and apply our model for computer simulation analysis in molecular dynamics. In our novel model, we propose to use implicit function in spatio-temporal predicate implementation. This allows us to pose complex shape queries changing over time. Our extended model allows us to integrate visual mining and querying of data in spatio-temporal databases or in files. Based on the proposed geometric model, we developed graphical user interface to visualize spatio-temporal data and to pose time-dependent complex shape queries on spatio-temporal data. We introduce three types of queries for data analysis in molecular dynamics and discuss possible hypothesis that can be formulated and tested using our system.

The proposed model has been tested in analysis of the results of molecular dynamics simulations studies of DNA interactions with ubiquitous positively charged proteins (the histones) in an all-atom system including explicit solvent (water) and ions. In our tests, the proposed Molecular Dynamics Visual Query System (MDVQS) was coupled to gOpenMol [8] program to provide the user with full specter of other analysis tools as well.

The paper is organized as follows. In the next Section, the novel geometric model for querying of time-dependent data is introduced. Then, an implementation of the model as a visual query system MDVQS is described. Query types for molecular dynamics system analysis are proposed. Finally, we discuss possible geometric hypothesis that can be formulated and tested using geometric queries. Conclusion and future work are presented as well.

GEOMETRIC MODEL FOR VISUAL MINING AND QUERYING OF SPATIO-TEMPORAL DATA

Let us introduce the formal mathematical specification of the geometric model for querying and mining of time-dependent data with the function-based representation of geometric solids. A geometric object can be a set of points

$$P = \{[x_1, x_2, \dots, x_n, t]\} = \{[\mathbf{X}, t]\}$$

in n dimensional Euclidean space E^n , and t is time. Primitive solid objects are defined with implicit functions as

$$f(x_1, x_2, \dots, x_n) \geq 0 \text{ in Euclidean space } E^n.$$

The implicit function $f(x_1, x_2, \dots, x_n) \geq 0$ can be defined analytically or by procedure. Such functions define closed n -dimensional objects in E^n space under the following conditions:

- $f(\mathbf{X}) > 0$ - for the points inside the object,
- $f(\mathbf{X}) = 0$ - for the points on the object boundary,
- $f(\mathbf{X}) < 0$ - for the points outside the object.

In our model, query solid can have time-dependent parameters and/or coordinates that can be defined analytically or by procedure. Thus, the geometric query model consists of the following geometric objects:

- n -dimensional points $P = \{[x_1, x_2, \dots, x_n, t]\}$ where t is time;
- time-dependent 3-dimensional primitive geometric objects for the construction of a query solid using geometric operations.

The following is an implicit function representation of the primitive time-dependent 3-dimensional geometric solids that could be used for construction of geometric criteria:

Halfspace:

$$G_1: f_1(\mathbf{X}, t) = f_1(x_1, x_2, x_3, t) = (x_1 - a[t]) \geq 0$$

where a is some real number ($a \in \mathbb{R}$)

Sphere:

$$G_1: f_1(\mathbf{X}, t) = r[t]^2 - (x_1 - x_{0,1}[t])^2 - (x_2 - x_{0,2}[t])^2 - (x_3 - x_{0,3}[t])^2 \geq 0$$

where $x_{0,1}, x_{0,2}, x_{0,3} \in \mathbb{R}$

Ellipsoid:

$$G_1: f_1(\mathbf{X}, t) = 1 - ((x_1 - x_{0,1}[t])/a_1[t])^2 - ((x_2 - x_{0,2}[t])/a_2[t])^2 - ((x_3 - x_{0,3}[t])/a_3[t])^2 \geq 0$$

where $x_{0,1}, x_{0,2}, x_{0,3} \in \mathbb{R}$ and $a_1, a_2, a_3 \in \mathbb{R}$.

Cone:

$$G_1: f_1(\mathbf{X}, t) = ((x_1 - x_{0,1}[t])/a_1[t])^2 - ((x_2 - x_{0,2}[t])/a_2[t])^2 - ((x_3 - x_{0,3}[t])/a_3[t])^2 \geq 0$$

where $x_{0,1}, x_{0,2}, x_{0,3} \in \mathbb{R}$ and $a_1, a_2, a_3 \in \mathbb{R}$.

Cylinder:

$$G_1: f_1(\mathbf{X}, t) = ((x_1 - x_{0,1}[t])/a_1[t])^2 - ((x_2 - x_{0,2}[t])/a_2[t])^2 \geq 0$$

where $x_{0,1}, x_{0,2} \in \mathbb{R}$ and $a_1, a_2 \in \mathbb{R}$.

By further declaring that our model is open to any type of objects that can be defined implicitly with some functions $f(x_1, x_2, x_3, t) \geq 0$, we could avoid the problem of a minimum set of primitives and to change this set depending on the application problem to be solved.

Geometric operations are applied to primitive geometric objects to obtain complex geometric shapes at each time point. The analytical definition of set-theoretic operations is realized in the form proposed by [13], where operations over implicit functions are considered. Affine transformations (translation, rotation and scaling) are also used to increase an expressive power of the proposed geometric model. Geometric operations include set-theoretic union, intersection, difference, and orthographic projection.

Mathematically,

Union: $G_3 = G_1 \cup G_2$ of two objects $G_1 \subset E^n$ and $G_2 \subset E^n$ with the descriptive functions f_1 and f_2 will be defined as $f_3 = f_1 \vee f_2 = \max(f_1, f_2) \geq 0$, where $G_3 \subset E^n$.

Intersection: $G_3 = G_1 \cap G_2$ of two objects $G_1 \subset E^n$ and $G_2 \subset E^n$ with the descriptive functions f_1 and f_2 will be defined as $f_3 = f_1 \wedge f_2 = \min(f_1, f_2) \geq 0$, where $G_3 \subset E^n$.

Complement: $G_2 = \neg G_1$ of object $G_1 \subset E^n$ with the descriptive functions f_1 will be defined as $f_2 = -f_1 \geq 0$.

Difference: $G_3 = G_1 \setminus G_2$ between objects $G_1 \subset E^n$ and $G_2 \subset E^n$ with descriptive functions f_1 and f_2 will be defined as $f_3 = f_1 \wedge (-f_2) = \min(f_1, -f_2) \geq 0$, where $G_3 \subset E^n$.

Translation: $G_2=T(G_1)$ of object $G_1 \subset E^k$ with descriptive functions f_1 by a_1, a_2, \dots, a_n will be defined as $f_1(x_1-a_1, x_2-a_2, \dots, x_n-a_n) \geq 0$.

Rotation: $G_2=R(G_1)$ of object $G_1 \subset E^k$ with descriptive functions f_1 of angle α about some axis will be defined as $f_1(x'_1, x'_2, \dots, x'_n) \geq 0$ where $[x'_1 \ x'_2 \ \dots \ x'_n \ 1] = R^{-1}[x_1 \ x_2 \ \dots \ x_n \ 1]$ and R^{-1} is an inverse matrix of rotation.

Scaling: $G_2=S(G_1)$ of object $G_1 \subset E^k$ with descriptive functions f_1 in s_1, s_2, \dots, s_n times will be defined as $f_1(x_1/s_1, x_2/s_2, \dots, x_n/s_n) \geq 0$.

We introduce a spatio-temporal point/solid predicate for query implementation. Let P be a point in Euclidean space E^n and t is time, G_1 be a query solid described with implicit function f_1 defined with time-dependent parameters and location changing over time, bG_1 be a boundary of G_1 and iG_1 be an interior of G_1 . Then a point/solid predicate is described with the implicit function representation of the geometric object G_1 by a 3-valued predicate:

$$S_3(P, G_1) = \begin{cases} 0, & \text{if } f_1(x_1, x_2, \dots, x_n, t) < 0 \quad P \notin G_1 \\ 1, & \text{if } f_1(x_1, x_2, \dots, x_n, t) = 0 \quad P \in bG_1 \\ 2, & \text{if } f_1(x_1, x_2, \dots, x_n, t) > 0 \quad P \in iG_1 \end{cases}$$

VISUAL MINING AND QUERYING

As we mentioned before, with our novel geometric model we can not only compute and classify atom trajectories inside or across the arbitrary shapes but we can also use computer graphics techniques to visualize and query spatio-temporal data interactively. We implemented Molecular Dynamics Visual Query System (MDVQS) with the graphical user interface GUI. The proposed geometric model for querying time-dependent data is meant to be the user's model as well as the formal foundation for the implementation of MDVQS. The points mapped from the spatio-temporal database, the constructed geometric solids, the query formulation, and the resulting geometric objects changing over time are visualized. To get an initial impression of data, the data are visualized as 3-dimensional snapshots of point clouds or molecular system at each time point. After the data is visualized as 3-dimensional projection at some time point, a complex geometric query solid with union, intersection or other operations over primitive geometric solids is posed by the user. These primitive query solids currently are cuboid (box), ellipsoid, cone, and cylinder. In addition, any point of clouds can be located and identified in the database or file as corresponding atom/molecules at time point. First, 3-dimensional projection of points/atoms/molecules is mapped from the

database/file and visualized as point clouds or molecular system. A blobby solid can be reconstructed at each time point to show shape of point clouds changing over time. Then, a solid query is posed and time interval is chosen. For instance, a query solid is a cylinder that does not change its parameters and location over time interval. The result of the query is time-dependent data and is visualized as set of snapshots or as a file with animation.

The developed geometric query tools allow the user to formulate spatio-temporal queries that cannot be implemented directly using any available spatio-temporal database system or molecular visualization program. With the proposed query model, the user could specify a query solid for each time point defining time-dependent primitive solids parameters and location analytically or by procedure. Currently, with the implemented GUI, the user constructs the query shape that does not change over time interval.

Geometric objects can be drawn opaque or transparent. We employ visualization techniques and advanced computer graphics algorithms for the implementation of the user interface. For better immersion stereo visualization is developed. The system is implemented in MS Visual C++ 6.0 with computer graphics software Visualization ToolKit (VTK) [14] where visualization is implemented with marching cube algorithm. VTK is an open source, freely available software system for 3D computer graphics, and image processing. We use VTK 4.2 C++ class library. MDVQS was tested on PCs (Windows NT/2000/XP).

MOLECULAR DYNAMICS APPLICATION

Based on the proposed geometric model, we developed the visual query system MDVQS for visual mining and querying results of numerical simulation in molecular dynamics. We work with the resulting data of numerical simulation studying structure and dynamics of DNA in interaction with solvent (water) different ions and positively charged fragments of the proteins. Data containing information about motion of all atoms in the system over the simulation time were generated by using CHARMM27 force field [3, 4]. Molecular dynamics algorithm apply classical Newtonian equations of motion and with self-consistent approximation for inter- and intramolecular forces (empirical potentials for chemical bonds, angles and torsions, electrostatic potential for ions and partially charged atoms, Lennard-Jones potentials for disperse interactions). With this numerical model, we are able to obtain data changing over time that we use as the input to our visual query system.

The time-dependent data consists of atom coordinates changing over time. Values are calculated at typically equal time intervals. Here, we use a common molecular dynamics “Xmol” format for input and output in our visual query system. The file format example is shown in Figure 1. The number in the first line indicates the number of atoms and molecules. The second

line specifies the snapshot time point. Other lines start with chemical symbols of the atoms or molecules followed by x , y and z coordinates of the atoms. There are many snapshots of data corresponding to time sequence. The MDVQS automatically visualize molecular system performing the molecular structure presentation.

The visual query system allows the user immediate verification of various hypotheses concerning the behavior of the DNA, the protein and their interactions. Realistic 3D visualization of the time sequence of data snapshots helps to detect unexpected behavior.

After studying the problems of molecular dynamics, we proposed the following three types of queries that can be easily implemented with our geometric model. Let us describe the queries.

Query Type 1. *Find and display trajectories of atoms by atom name or by its location (x, y, z)*

The user will be able to pick specific atoms that he/she is interested in, by their name or by location, and we will display only those selected atoms changing its location over time. For both parts, the number of atoms that could be chosen would be limited due to the large amount of data that the system may go through. Our tests were done with 500 files of 40 MB each.

Query Type 2. *Find and display snapshots of atoms over time in the selected region where the region is the final query solid constructed as a result of operations over primitive solids.*

The region can be of various shapes such as a box, an ellipsoid, a cylinder, a sphere or even a cone. The result of union, intersection, and/or subtraction operation over the primitive solids can be a query solid as well.

Query Type 3. *Find trajectories of atoms for a specified time interval $[t_1, t_2]$*

We can combine all three queries in different combinations. This will allow the user to choose specific atoms and at the same time an interval and a region in which the atoms moves over time. This will be particularly useful for observing when the chosen atoms move out of the region specified along their trajectory.

Let us show some application examples of querying process with our tools. The amount of MD data describing dynamic behavior with all-atom precision for realistic nanoscale system (tens to hundreds of thousands atoms simulated for time inter-

vals approaching microseconds) with 0.1-10 picosecond time interval (to catch fast movements of water molecules and other mobile components) may exceed hundreds of Gigabytes. In our example, the system consists of 34,155 atoms: three fragments of the DNA double helix (22 base pairs in each fragment) in hexagonal simulation cell with addition of 14 similar oligopeptides (each of 8 amino acids), 84 K^+ ions, and 9430 water molecules (snapshot of the system visualized with MDVQS is shown in Fig. 2). The system was simulated for more than 25 nanosecond; coordinates of the atoms were recorded each 0.1 picosecond. Usage of visual mining and querying allows separating defined parts of the system in order to observe certain atom-atom interactions and molecule transformations occurring in the system which otherwise (without application of visual mining and querying tools) maybe missed in the “noise” of numerous other events. In Figure 2, DNA molecules and histone protein fragments displayed in ball-stick presentation, potassium ions are drawn as magenta spheres, water molecules not shown.

With MDVQS we can pose a query in two ways: visually, by drawing the shapes or by entering parameters of shapes. All parameters measures are in Å. If we pose the query by entering parameters of shapes it can be formulated as follows “find all atoms/molecules within the ellipsoid with semi-axes $a_1=6$ Å, $a_2=6$ Å, $a_3=2$ Å.” In Figure 3, the result of tube solid query per time point is shown. The tube solid query is posed visually at time point where the query solid is the difference between two cylinders with radiuses r_1 and r_2 correspondently, where $r_1 < r_2$. In the molecular system modeling the behavior of the double helical DNA such a query allows to highlight interaction of the negatively charged phosphate group (yellow-red groups of atoms in figure) with protein, K^+ and water. Also flexibility of the DNA is tested by visualization of the DNA atoms entering and leaving the defined space. In Figure 4, the result of querying with the solid defined as an ellipsoid per time point is shown. In the exemplified system such a query allows observation of the diffusion of the atoms from DNA and other components through particular region of the simulation cell. In Figure 5, the result of querying with the solid defined as an intersection of four halfspaces is shown. This query is applied to separate motion of several DNA base pairs from the other events in the system. In Figure 6, the result of querying with the solid defined as a cone is shown. This query indicates binding of the protein, K^+ and other species to a phosphate group. This group is situated on the surface of the DNA double helix and the space accessible for the mobile species near the phosphate has a conical shape.

With MDVQS the query result is visualized as set of snapshots or as a file with animation. Thus, posing the queries help the user to analyze molecular structures and their physical and chemical properties. The user is interested in the interaction be-

tween certain atoms in the molecular structure, and the queries allow him/her to be able to focus the attention on the motion of certain atoms.

The results of the queries can be visualized with the system gOpenMol Version 2.32 as well. The user can go through the time point snapshots of the resulting molecular structure to take advantages of using all functions available in gOpenMol. System gOpenMol is a tool for visualization and analysis of molecular structures combined with several applications for data analysis and presentation originated from quantum mechanics, molecular dynamics and other computational chemistry calculations. The system can be downloaded from the Internet (<http://www.csc.fi/gopenmol>). We also can view results of the queries as the trajectories of molecules using gOpenMol software. The queries implemented in MDVQS complement the functions of gOpenMol system. MDVQS works with gOpenMol system through common “Xmol” file format. gOpenMol can be invoked from MDVQS interface.

The developed visual query system MDVQS can be easily extended to access other formats that are used in other popular molecular visualization programs.

We proposed the queries for data analysis in molecular dynamics that allow the user formulate new hypotheses. Let us describe possible geometric hypotheses that can be tested using the proposed queries.

Hypothesis 1. *The molecular structure does not change and/or does not move considerably over time.*

The user picks atoms/molecules by its name or by location, and predicts the area of its location in space. The region can be of various shapes such as a box, an ellipsoid, a cylinder, a sphere or even a cone.

Hypothesis 2. *The chosen atoms and/or part of the molecular structure are always located inside the fixed corridor.*

The user defines the predicted corridor. Then, the program will test not only the validity of the hypothesis for each time frame but also will show which atoms intersect the borders of the corridor. For example, this type of hypothesis can be formulated by querying the time frames with geometric shape that is the result of difference of two cylinders with radiuses r_1 and r_2 , where $r_1 < r_2$, or with geometric shape that is the result of difference of two ellipsoids with semi-axes a_1, b_1, c_1 and a_2, b_2, c_1 where $a_1 < a_2, b_1 < b_2$, and $c_1 < c_2$. Observation of the frequency and identities of the atoms invading the corridor is also valuable.

Hypothesis 3. *The corridor in which the chosen atoms are located changes its shape over time.*

The user defines the predicted corridor as geometric shape changing over time.

The list of hypotheses can be extended by adding geometric shapes into computational and/or visualization system according to the needs of user and specificity of the system under study. The only condition is the user's ability to describe the new shape with implicit function representation. We also can introduce more complicated hypotheses by constructing arbitrary shape with the different geometric operations.

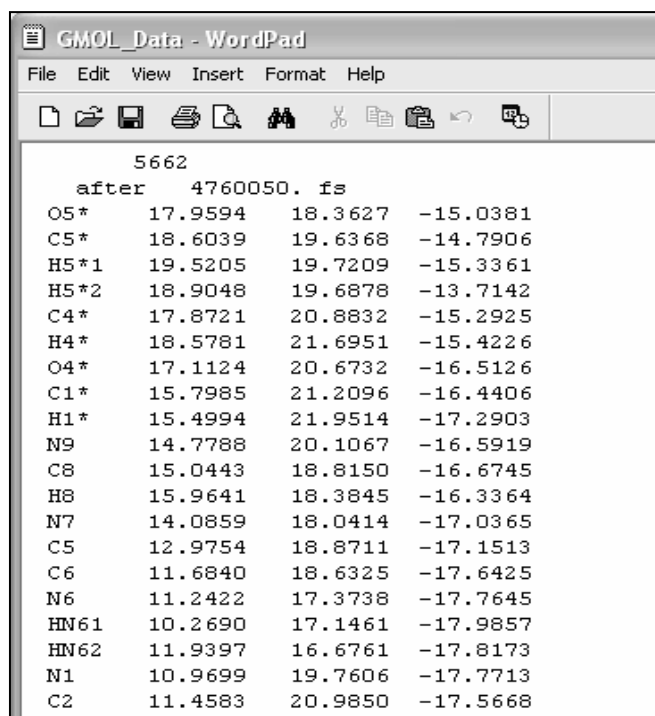
The geometric queries in our model can be considered as motion classifiers. The resulting sets are coordinates of molecules and atoms inside the particular shape changing over time.

CONCLUSION AND FUTURE WORK

In this paper, we introduced a novel geometric model that allows us to test validity of different hypotheses about evolution of molecular systems using different geometrical shapes and their combinations. We proposed query types for data analysis in molecular dynamics. The proposed geometric model and developed geometric query tools are unique and allow the user to formulate spatio-temporal queries that cannot be implemented directly using any available spatio-temporal database system and/or molecular visualization program. Based on the uniform geometric model we integrate visual mining with querying of time-dependent data in one GUI. The user can come up with different geometric hypotheses and test their validity using the system. In our model, we also proposed solid queries with time-dependent parameters and location that can find its further application in domain of molecular dynamics or other potential application areas where time-dependent data are studied. The proposed query model can be also easily extended to query time-dependent solids. Now, we are working on time-dependent data representation model to allow implementation of efficient spatial access methods and structures. We are working on improving the user interface of MDVQS making it more user-friendly as well. The proposed innovative approach could be especially useful in analysis of the molecular dynamics data obtained for objects of very large size typical for experimental nanoscience. In such big systems, querying gives possibility to separate specific motions and transformations from myriads of other events simultaneously occurring in the simulation cell while keeping all the MD data intact for the other sorts of analysis.

REFERENCES

1. D. A. Keim, Information Visualization and Visual Data Mining, in IEEE Trans Vis and Comput Graph (2002), Vol. 8(1), pp.1-8
2. M. H. Dunham, Data Mining Introductory and Advanced Topics, Person Education, New Jersey (2003)
3. N. Foloppe and A. D. MacKerell, All-Atom Empirical Force Field for Nucleic Acids: 2) Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data, in J.Comp.Chem. (2000), Vol. 21, pp.86-104
4. A. D. MacKerell and N. Banavali, All-Atom Empirical Force Field for Nucleic Acids: 2) Application to Molecular Dynamics Simulations of DNA and RNA in Solution, in J.Comp.Chem. (2000), Vol. 21, pp.105-120
5. J. W. Ponder and D. A. Case, Force Fields for Protein Simulations, in Adv.Protein Chem. (2003), Vol. 66, pp.27-85
6. E. Lindahl, B. Hess and D. van der Spoel, Gromacs 3.0: A Package for Molecular Simulations and Trajectory Analysis, in J.Mol.Mod. (2001), Vol. 7, pp.306-317
7. W. Humphrey, A. Dalke and K. Schulten, VMD - Visual Molecular Dynamics, in J.Mol.Graph. (1996), Vol. 14, pp.33-38
8. D. L. Bergman, A. Laaksonen and L. Laaksonen, Visualization of Solvation Structures in Liquid Mixtures, in J.Mol.Graph. Model. (1997), Vol. 15, pp.301-306
9. R. H. Güting, M. H. Böhlen, M. Erwig, C. S. Jensen, N. A. Lorentzos, M. Schneider and M. Vazirgiannis, A Foundation for Representing and Querying Moving Objects, in ACM Trans on Database Systems (2000), Vol. 25(1), pp.1-42
10. M. Erwig and M. Schneider, Developments in Spatio-Temporal Query Languages, in IEEE Int. Workshop on Spatio-Temporal Data Models and Languages (STDML) (1999), pp.441-449
11. O. Sourina and S. H. Boey, Geometric Query Types for Data Retrieval in Relational Databases, in Data & Knowledge Engineering, Elsevier Science B.V. (1998), Vol. 27(2), pp.207-229
12. J. Bloomenthal, An Introduction to Implicit Surfaces Morgan-Kaufmann, San Francisco (1997)
13. A. Ricci, A Constructive Geometry for Computer Graphics, in The Computer Journal (1973), Vol. 16(2), pp.157-160
14. W. Schroeder, K. Martin and B. Loresen, The Visualization Toolkit An Object-Oriented Approach To 3D Graphics 3rd Edition (2000), <http://www.kitware.com/products/vtktextbook.html>.



The image shows a screenshot of a WordPad window titled "GMOL_Data - WordPad". The window contains a table of coordinates in Xmol format. The table has four columns: an atom label, and three numerical values representing coordinates. The data is as follows:

Atom Label	X	Y	Z
5662			
after 4760050.fs			
O5*	17.9594	18.3627	-15.0381
C5*	18.6039	19.6368	-14.7906
H5*1	19.5205	19.7209	-15.3361
H5*2	18.9048	19.6878	-13.7142
C4*	17.8721	20.8832	-15.2925
H4*	18.5781	21.6951	-15.4226
O4*	17.1124	20.6732	-16.5126
C1*	15.7985	21.2096	-16.4406
H1*	15.4994	21.9514	-17.2903
N9	14.7788	20.1067	-16.5919
C8	15.0443	18.8150	-16.6745
H8	15.9641	18.3845	-16.3364
N7	14.0859	18.0414	-17.0365
C5	12.9754	18.8711	-17.1513
C6	11.6840	18.6325	-17.6425
N6	11.2422	17.3738	-17.7645
HN61	10.2690	17.1461	-17.9857
HN62	11.9397	16.6761	-17.8173
N1	10.9699	19.7606	-17.7713
C2	11.4583	20.9850	-17.5668

Fig. 1. Example of file in "Xmol" format.

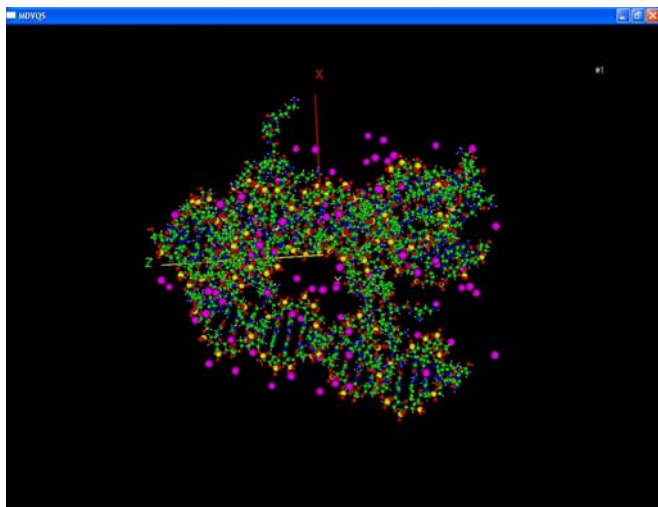


Fig. 2. Molecular system per one time frame.

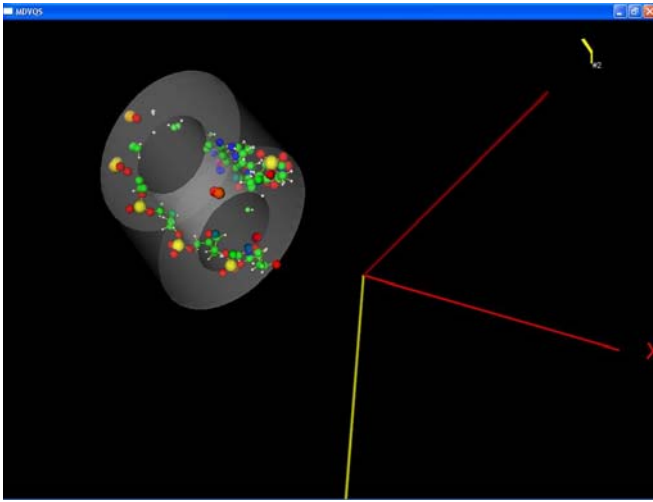


Fig. 3. Querying with the solid defined as a difference of two cylinders.

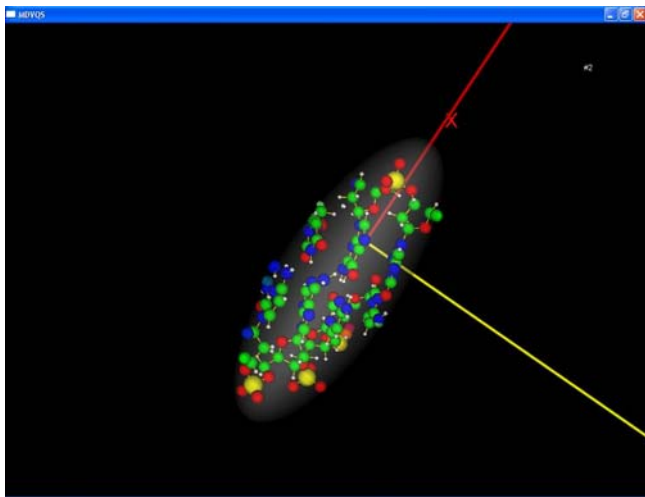


Fig. 4. Result of querying with the solid defined as an ellipsoid.

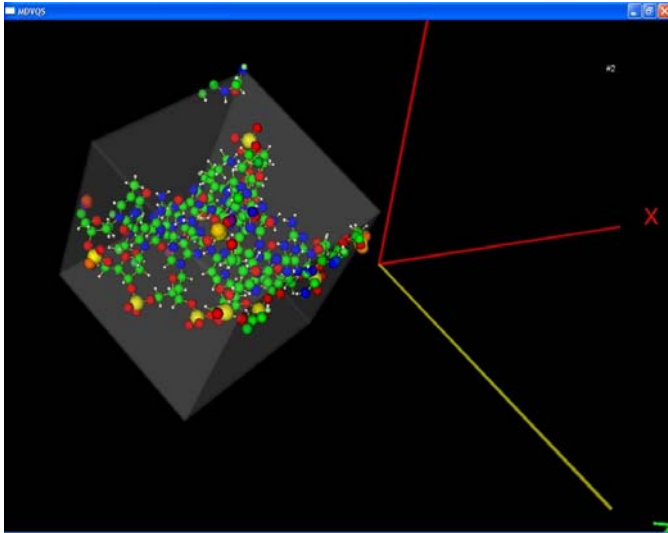


Fig. 5. Result of querying with the solid defined as an intersection of four halfspaces.

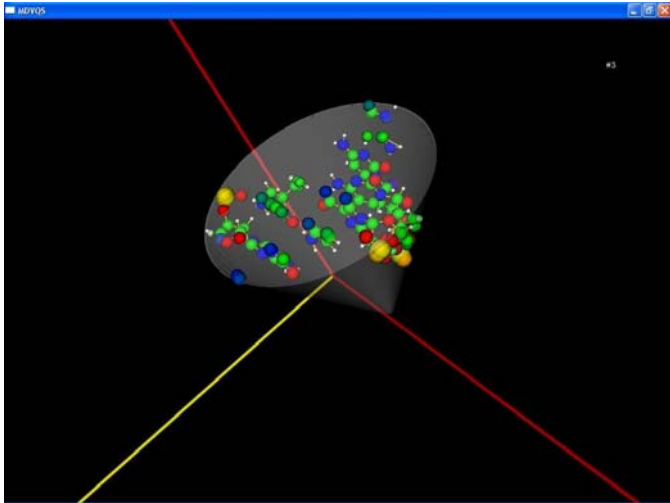


Fig. 6. Result of querying with the solid defined as a cone.