

Detecting Events in Molecular Dynamics Simulations

Iris Adä and Michael R. Berthold

Nycomed-Chair for Bioinformatics and Information Mining
Dept. of Computer and Information Science
University of Konstanz
`first.last@uni-konstanz.de`

Abstract. We describe the application of a recently published general event detection framework, called *EVE* to the challenging task of molecular event detection, that is, the automatic detection of structural changes of a molecule over time. Different types of molecular events can be of interest which have, in the past, been addressed by specialized methods. The framework used here allows different types of molecular events to be systematically investigated. In this paper, we summarize existing molecular event detection methods and demonstrate how *EVE* can be configured for a number of molecular event types.

1 Introduction

Research in Chemistry/Chemical Biology has been interested in understanding molecular dynamics simulations for around 25 years [14,19]. There are many different ways of modeling molecules; here we are primarily interested in molecules (as in Figure 1) represented as a connected group of multiple atoms, ranging

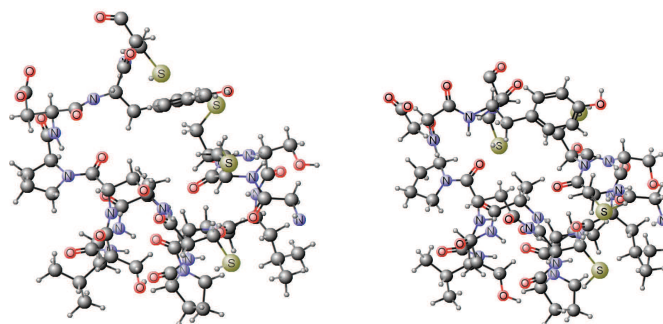


Fig. 1. Two conformations of alpha-conotoxin pnib (1AKG). The molecule's atoms are connected by bonds and the entire arrangement in space represents the molecule's conformation.

from tens to several hundreds of atoms and more. The geometrical arrangement of atoms to each other (known as the *conformation* of the molecule) under normal conditions changes continuously, even in a vacuum. These changes range from small oscillations to drastic changes in the overall molecular shape. It is important to differentiate between small, unimportant conformational changes and more important ones, which are relevant e.g. for the biological function of a molecule. The automatic detection of unexpected or irregular transformations of the molecule's conformation is of particular interest.

There are two main reasons why molecular dynamics (MD) is concerned with simulating data. First of all, monitoring atom positions and hence the molecular conformation at a sufficiently high resolution (both in terms of time and location) is complicated and second the influence of the molecule's surroundings has a substantial effect on the molecule's conformations. Using simulated movements over time helps to better understand the behavior of a molecule. However, actually processing the vast amount of data generated by these simulations poses enormous problems.

The core idea behind molecular dynamics simulations is to simulate the behavior of molecules, mostly of proteins, over time by using "simple potential-energy functions" [2]. Hence this is an artificially, but not randomly generated data set, which depicts the true behavior of the molecule. The forces and reciprocal effects between the molecule's atoms and bonds can be explained with fairly simple mathematical functions. Basically every atom can influence all other atoms with the bonds between the atoms adding to this effect. The dynamics are calculated by solving equation systems based on all these functions together [2]. By iteratively calculating the steps of the molecule's internal positions a sequence of conformations is generated. This sequence captures the movement of the molecule, represented by all of the individual atom positions in three dimensional coordinates over time. The interest of molecular event detection lies in finding unexpected movements among atom positions of a molecule. Such movements can e.g. refer to a conformational change or a folding of the molecule.

The internal relations between molecule's atoms and bonds are investigated for other aspects (e.g. kinetic and thermodynamic information [2]) as well in molecular mechanics [8]. However, in this work we are interested only in the changes of atom positions over time, hence, in the sequence of molecule conformations. There are various types of molecular events. The key point of interest is that these structural changes are relevant for the chemical state or mechanism of the molecule.

In this work we concentrate on the problem of general molecular event detection as an application for change and event detection in high dimension. We begin by discussing current approaches, starting with an overview of feature based methods before discussing more recent methods. Afterwards we briefly summarize the concepts of **EVE** [1], an event detection framework we use to formalize the underlying event detection problem. We conclude by demonstrating how **EVE** can be used to detect molecular events of interest for a real molecular sequence.

2 Related Work

In the area of data mining and statistics considerable work has already been invested in the detection of irregularities in series data. Event detection [12], drift detection [24] and anomaly detection [9] are vibrant research directions.

In this section we focus on how events are detected in molecular dynamics. One of the main difficulties is the high dimensionality of the feature vector. Following the structure of a recent survey paper [6] we will first discuss traditional analysis methods, which are mainly concerned with the extraction of features. Recent methods regard the molecule’s atoms as the nodes in a graph and apply different strategies to introduce edges and then monitor changes of the resulting graph over time.

2.1 Traditional Analysis

A large number of atoms – easily hundreds – is encountered when analyzing a molecular times series. Using their coordinates and other properties as one huge feature vector is time consuming and often yields uninterpretable results. For this reason quite a few applications of molecular event detection are first concerned with preprocessing the features of the molecule’s conformation.

H Atom Filtering: As already mentioned, molecules consist of many atoms like carbonate (C), hydrogen (H), or oxygen (O), to name just some of the common ones. Smaller atoms (like hydrogen), for example, are known to be more prone to move. Common practice therefore ignores the movements of hydrogens as they are not related to an overall structural change of the molecule and can easily be derived from the remainder of the molecular structure anyway.

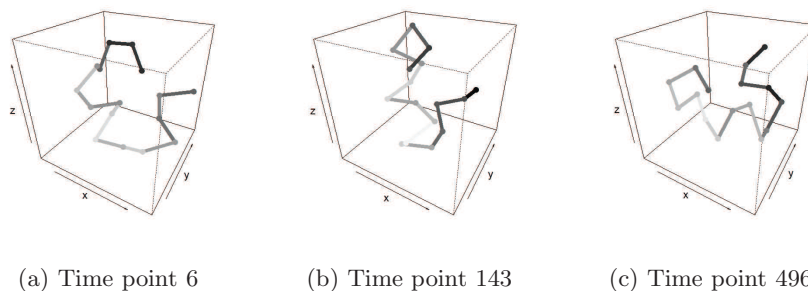


Fig. 2. The C^α trace of alpha-conotoxin pnib. The C^α atoms are connected to show the overall structure of the molecule. This reduces the complexity tremendously, as only 16 of the 213 atoms in the molecule remain. The figures show three conformational states of the molecule, at time points 6, 143 and 496. At step 143 the molecule is folded and opened again in a last step (496). The three time points were chosen because they demonstrate the movement of the molecule nicely.

C^α Atom Extraction: A more compact representation retains only the so called C^α trace [11], where a C^α -carbon is the central carbon of an amino acid [21]. Put simply, each protein consists of multiple amino acids, which are “substructures” of the molecule. In each amino acid the C^α -carbon can be uniquely determined and all of these carbons are combined to the C^α trace, which is an abstraction of the overall structural appearance. Therefore only the C^α carbons are used as features and all other atoms are filtered. This reduction naturally reduces the dimension immensely. It is questionable whether it also filters a possible movement from the data. In Figure 2 the C^α trace of 1AKG is visualized for three consecutive time points.

Distances: One problem when monitoring molecular dynamics is that only relations within the molecule’s atoms are of interest and not an overall movement. One solution for this problem is to calculate relative distances (e.g. Euclidean or absolute) between each pair of atoms inside the molecule. The resulting feature space is quadratic in the number of atoms.

Angles: A second solution for filtering an overall movement from the data is to use angles between the position of individual atoms. The feature vector then consists of the angles in three dimensional space.

RMSD: A “well-known and most widely accepted” [15] method for measuring the similarity between two conformations of a molecule is the root-mean-square deviation (RMSD). It is calculated by using the average distance between all atom positions.

$$\text{RMSD} \left(M^{(1)}, M^{(2)} \right) = \sqrt{\frac{1}{n} \sum_{i=1}^n |\mathbf{m}_i^{(1)} - \mathbf{m}_i^{(2)}|^2}$$

$M^{(t)}$ is the conformation of molecule M at time point t and $\mathbf{m}_i^{(t)} \in \mathbb{R}^3$ is the atom position of atom i .

The disadvantage is that the RMSD measure also reacts to movements and rotations of the molecule. The problem of movements can be resolved by shifting one molecule and rotation can be solved by pre-applying a rotation. The Kabsch Algorithm [13] is one of the most popular solutions for this optimization problem and uses a singular value decomposition to find the minimizing transformation matrix. A more recent approach uses Quaternions to solve the problem [10].

Visualization: The detection of an event can also be determined by visual inspection. In a line plot of the RMSD measure, molecular events are mapped to peaks in the dissimilarity. More recently heatmaps of the complete distance matrix between all time points are used as well. Areas showing a small in-between dissimilarity and a large dissimilarity to neighboring times are then further investigated.

2.2 Non-traditional Analysis

During the last few years, new methods for analyzing MD simulations came up. In contrast to traditional analysis methods they are designed to deal with much larger simulations.

In flexibility analysis [22,4] each atom is individually investigated using a principal component analysis (PCA). The results enable atoms with fast vibrations or small movements to be filtered and summarize the major states of the molecule. Finally the flexibility vectors, generated from the PCA, are plotted on a mean structure of the molecule for further analysis [6]. In the wavelet analysis of MD simulations [3,7] the atoms are individually analyzed as well. Using a continuous wavelet transform Benson and Daggett [7] are able to find trajectories of different proteins that show similar structural movements.

The most recent research direction uses graphs to model the overall dependencies in the molecule. Wriggers et al. [23] presented probably one of the earliest approaches. In this work a graph is generated from the atoms' positions. They propose multiple methods for generating the graph using a distance cut-off, whereby atoms are connected in the graph if their distance is below a certain threshold, or the generalized masked delaunay tetrahedralization. Afterwards changes in the consequent graphs are tracked over time, enabling the number of appearing and disappearing edges to be counted. Finally they apply different filters to the achieved series of graph changes to detect the event. The Dynamic Tensor Analysis [17,18] applies tensor analysis to identify conformational substates of the molecular sequence. More recently graphs are generated using additional expert knowledge of the chemical structures [5].

3 Goals of Molecular Event Detection

As described previously molecular event detection is concerned with the analysis of movements of molecules over time. The molecule's movement is simulated in three dimensional space and every few picoseconds, or even less, a snapshot/conformation of the current positions is calculated. The molecule M is represented as list of its three dimensional atom positions $M = (\mathbf{m}_1, \dots, \mathbf{m}_n)^t$, where $\mathbf{m}_i = (x, y, z)^t \in \mathbb{R}^3$. The molecules are generated at time points $t = 1 \dots m$. The molecule at time point t is entitled $M^{(t)}$.

The goal of molecular event detection is to find interesting changes in the consequent states of a molecule. Next we assume that a dissimilarity function $d(\cdot, \cdot)$ is provided, which is able to calculate the dissimilarity between two conformations of a molecule. For example RMSD can be used.

Constant: A molecular dynamics is called constant in a time interval $[t_0, t_1]$ if the relations in-between the atom positions do not change
 $(\exists \epsilon \geq 0 \forall i, j \in [t_0, t_1] : |d(M^{(i)}, M^{(j)})| < \epsilon)$.

Changing: Molecular dynamics change if the dissimilarities increase over consequent time steps. A molecule is changing in the time interval $[t_0, t_p] = (t_0, \dots, t_p)$

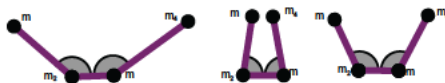


Fig. 3. The process of molecular folding/refolding, illustrated by showing four exemplary significant atoms during the process

if the dissimilarity to predecessors increases, hence
 $\forall i \in \{1, \dots, p-1\} : d(M^{(t_0)}, M^{(t_i)}) < d(M^{(0)}, M^{(t_p)})$.

Reoccurring Change: A change is regarded as reoccurring if the atoms return to their previous relations. With respect to the definition of changing, there would be a second interval where the dissimilarities to predecessors decrease. The interval $[t_0, t_1]$ is called a reoccurring change if $\exists t_2, t_3 : [t_2, t_0] \cup [t_1, t_3]$ is constant. Note that this type of change is mostly tripartite, starting with a change, followed by a constant state and finishing with the reoccurrence to the beginning structure.

Outlier: An outlier is defined as a single data point (or only a few) within a constant state where the molecule shows high dissimilarity to the previous and following steps. More formally: t is an outlier, if $\exists t_0, t_1, t \in [t_0, t_1] : [t_0, t_1] \setminus t$ not changing. Note that if the step size of the molecular simulation is fine enough, one would not expect to find such outliers.

A special kind of a reoccurring change is a folding and refolding of the molecule, which is a particularly interesting molecular behavior. During this event the molecule changes its shape into a new conformation (where it usually has a different biological function) and, after a certain time, refolds into the original conformation. Of key interest here is the identification of the exact start and end of the process. Figure 3 shows the visualization of a folding with subsequent refolding. We will now focus on illustrating below how these exemplary molecular event types can be detected using EVE.

4 EVE for the Detection of Molecular Events

This section presents an exemplary setup to demonstrate how EVE is configured for the application of molecular event detection. In the following the main aspects of the EVE framework are summarized: for a more detailed discussion see our previous work [1]. We consequently analyze 1AKG as a relatively small but well understood example of a molecule.

4.1 The EVE Framework

EVE [1] is a general framework for event detection. The main intention behind the framework is twofold. The first goal is to model the detection of events in a general setting. Using EVE it is straightforward to classify existing approaches as well as emphasizing dissimilarities and similarities between individual event detection techniques. Secondly, the framework is intended to provide a fast entry point to examine new, also complex data types. This fits perfectly for the application of molecular event detection.

EVE structures any event detection algorithm into three key components: window configuration, dissimilarity measure and detection mechanism.

1. *Window configuration*

Two windows need to be defined, the past and the current window. The past window models behavior which is assumed to be normal at this point in time. The current window on the other hand is tested to see if it contains a possible event. The process of a window over consequent time steps is defined by two terms. The first refers to the start position which can be sliding or fixed (short: S or F) and the second one to the window size which is either constant or growing (short: C or G). Two such windows, one for past and one for current (e.g. SC \leftrightarrow SC), determine the process of a window over consequent time steps as a *window combination*.

- FC \leftrightarrow SC (Fixed Constant to Sliding Constant): Comparing a window at the start of the series with a sliding window ending at the most recent data point.
- FC \leftrightarrow FG (Fixed Constant to Fixed Growing): Comparing the start window with the rest of the following data.
- FG \leftrightarrow SC (Fixed Growing to Sliding Constant): The idea behind this concept is to extract statistics or models out of all past information and compare it to the most recent data points.
- SC \leftrightarrow SC (Sliding Constant to Sliding Constant): Comparing two consequent sliding windows.

2. *Dissimilarity measure*

The goal of the dissimilarity function is to indicate the probability of an event being detected. By previously building an abstracted model on the window, multiple possibilities of calculating this dissimilarity are possible. The Kullback-Leibler distance [16], the Euclidean distance (L_2), other L-norms or classification measures (e.g. the false positive rate) are obvious examples.

3. *Detection mechanism*

Detection is performed to identify events by evaluating the previously calculated dissimilarity measures, often by simply applying a threshold function or using a control chart [20].

After introducing the main ingredients of EVE, we now demonstrate how it can be applied to the specific application of molecular dynamics and illustrate how relevant events of different type can be detected.

4.2 Experimental Setup

For the setup of the EVE framework the representation, dissimilarity function, and threshold are chosen as follows.

Data Model: When a window contains multiple conformations the positions of the atoms are averaged to filter out small oscillations of the atoms. This average structure is calculated as follows:

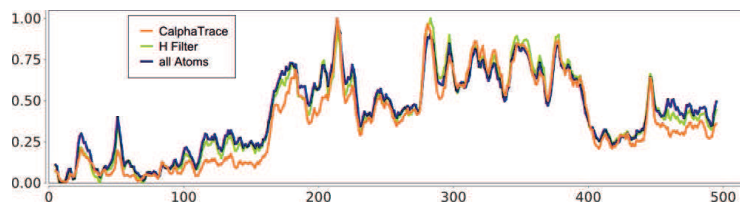


Fig. 4. The 0-1-normalized distance of three representation of one molecule (1AKG) with the same EVE setup ($\text{FG} \leftrightarrow \text{SC}$ and window size 5)
(**Legend:** x axis: time (ns), y axis: dissimilarity)

$$\text{avg} \left(M^{(1)}, \dots, M^{(c)} \right) = \frac{1}{c} \left(\sum_{i=1}^c m_1^{(i)}, \dots, \sum_{i=1}^c m_n^{(i)} \right).$$

The Root-mean-square deviation (as introduced in Section 2.1) is used as **dissimilarity measure**.

As **detection** mechanism a control chart [20] is used. An event is reported if the dissimilarity measure exceeds the sum of mean μ and 1.5 fold std. dev. σ .

We investigated four different window combinations:

$$\text{FC} \leftrightarrow \text{SC}, \text{FC} \leftrightarrow \text{FG}, \text{FG} \leftrightarrow \text{SC}, \text{SC} \leftrightarrow \text{SC}.$$

In the experiments we used three window sizes: $c \in \{1, 10, 20\}$.

We here previously filtered the H atoms of the molecule. In Figure 4 the $\text{FG} \leftrightarrow \text{SC}$ with window size 5 is applied on three representations, C^α -trace, H atoms filtering and all atoms. We normalized the distance measures to 0-1. However the visualization shows that the results are very similar and especially filtering the H atoms only had small effects to the overall error measure.

4.3 Molecular Data Set

The analysis is demonstrated on 1AKG, alpha-conotoxin pnib¹ (see Fig. 1), which is a relatively small molecule, actually a protein, containing 16 amino acids and 213 atoms. We chose this molecule because it is already well studied and hence the ground truth, e.g. the interesting events in the series are known.

The behavior of the molecule's conformation is as follows: During the first 70 time steps it does not change significantly, only small movements inside and a small overall rotation occur. Afterwards the molecule starts to fold. At time point ~ 230 folding reaches a maximum and the folding angles start to decrease again until time point ~ 400 , where a constant conformation is reached.

In Figure 6 to 8 three windowing concepts are applied to the alpha-conotoxin pnib over time. The x axis of each plot represents the time and the y axis the calculated dissimilarity. In all of these visualizations, the dissimilarity measure is shown in green, the lower bound of the control chart is depicted in blue and the upper bound is orange.

¹ Detailed information can be found in the protein data bank:
<http://www.rcsb.org/pdb/explore.do?structureId=1akg>

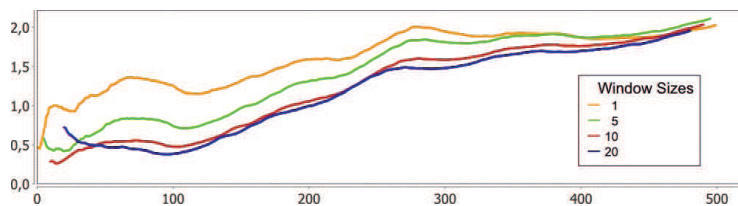


Fig. 5. $FC \leftrightarrow FG$ (Fixed Constant to Fixed Growing): When comparing a start window to the rest of the data there are two assumptions. First the past window has to represent the data nicely. The model of the current window should be able to smoothen events over time. This is not the case here, as the constant increase of all four windows does not reveal any insights.

(**Legend:** x axis: time (ns), y axis: dissimilarity)

4.4 Discussion of Experiments

In this section we discuss the analysis results of the molecular event detection using EVE.

The first type of window combination investigated is the $FC \leftrightarrow FG$ because it is the one with the most irrelevant result. Figure 5 shows the error calculated for the $FC \leftrightarrow FG$ on four different window sizes. The constant phase in the beginning of the molecular series is determined. However, afterwards the mean structure of the current window is not representative as it contains the change and the base line.

In addition to the changes of interest, the molecules' conformation changes steadily as the atoms are constantly see-sawing and there is always a lot of movement in smaller regions as well. This is clearly visible e.g. in the $SC \leftrightarrow SC$ of window size $c = 1$ (Figure 6a). There is so much movement in the dissimilarity function that the number of false positives is much too high. These oscillations can be filtered by adjusting the window sizes accordingly. However using a window size that is too big can smooth out possible events. In $FC \leftrightarrow FG$ (Figure 5)

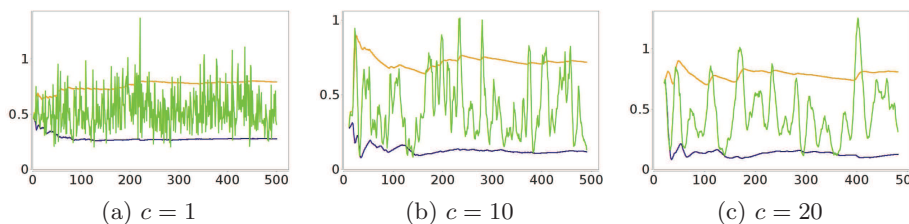


Fig. 6. $SC \leftrightarrow SC$ (Sliding Constant to Sliding Constant): Comparing two sliding windows worked out very nicely for 1AKG. While smaller window sizes contain much noise, $c = 20$ clearly show 3 very significant events.

(**Legend:** x axis: time (ns), y axis: dissimilarity, green: dissimilarity measured, blue: $\mu - 1.5 * \sigma$, orange: $\mu + 1.5 * \sigma$)

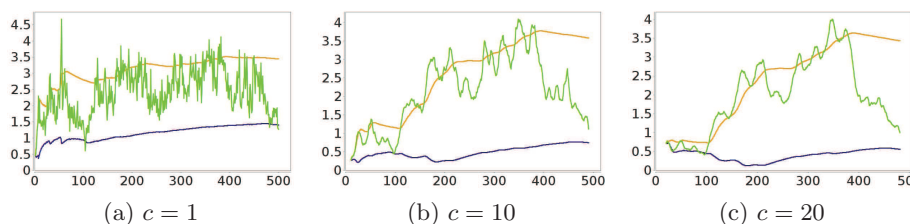


Fig. 7. $FC \leftrightarrow SC$ (Fixed Constant to Sliding Constant): In this analysis one can see, that the conformation of the molecule is less similar to the original state in the middle of the series and in the end returns to the initial conformation.

(**Legend:** x axis: time (ns), y axis: dissimilarity, **green:** dissimilarity measured, **blue:** $\mu - 1.5 * \sigma$, **orange:** $\mu + 1.5 * \sigma$)

an increasing line only is visible for all combinations, due to the fact that the second window is too big to produce meaningful results. However the opposite concept $FG \leftrightarrow SC$ (as can be seen in Figure 8 works well. The outlier is observed with the window size $c = 1$ (Fig. 8a).

The second part of our analysis was concerned with the question whether a reoccurring change can also be detected in the data. Two insights were provided by the $FC \leftrightarrow SC$ combination. The first part of the window up to time point 100 does not change very much, however, there is an outlier at time point 52 indicating a high dissimilarity to the baseline window. Although the outlier was smoothed out by bigger windows, they showed the overall movement trend in the conformational states of the molecule much more clearly. Considering the dissimilarity for $FC \leftrightarrow SC$ and window size $c = 20$, the constant phase in the beginning, the change, which was previously marked as a folding; and finally the return to the initial state can be seen clearly (Figure 7). The reoccurring change also becomes clear by inspecting Figure 8b where the constant behavior is reached again after time step 400.

The last observation is provided by the plots of the $SC \leftrightarrow SC$ combination (Figure 6). A window size of $c = 10$ yields good results on the first view. When

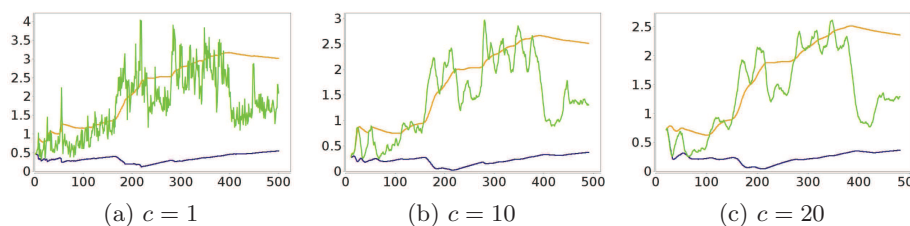


Fig. 8. $FG \leftrightarrow SC$ (Fixed Growing to Sliding Constant): The beginning and the end of the change can be seen in all three error plots. However the recoring can not be detected with the control chart.

(**Legend:** x axis: time (ns), y axis: dissimilarity, **green:** dissimilarity measured, **blue:** $\mu - 1.5 * \sigma$, **orange:** $\mu + 1.5 * \sigma$)

the default threshold ($\mu + 1.5 * \sigma$) is used, the start of the change, which should be detected at around time 70, is not detected. The first event would therefore not have been recognized. Using a window size of $c = 20$, on the other hand, does allow this event to be detected.

5 Conclusion

In this paper we demonstrated how the **EVE** framework can be applied to a complex problem, namely molecular event detection. Using this framework we were able to identify the points of interest and demonstrated that **EVE** can be used to detect events in challenging real world problems in a structure manner.

Acknowledgements. This research was partly supported by the DFG under grant GRK 1042 (Research Training Group “Explorative Analysis and Visualization of Large Information Spaces”).

We would like to thank Dr. Thomas Exner and Fredrick Robin Devadoss for providing the data and explaining many chemical details of molecular dynamics.

References

1. Adä, I., Berthold, M.R.: Eve: a framework for event detection. *Evolving Systems* 4(1), 61–70 (2013)
2. Adcock, S.A., McCammon, J.A.: Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical Reviews* 106(5), 1589 (2006)
3. Askar, A., Cetin, A.E., Rabitz, H.: Wavelet transform for analysis of molecular dynamics. *The Journal of Physical Chemistry* 100(49), 19165–19173 (1996)
4. Benson, N.C., Daggett, V.: Dynameomics: Large-scale assessment of native protein flexibility. *Protein Science* 17(12), 2038–2050 (2008)
5. Benson, N.C., Daggett, V.: A chemical group graph representation for efficient high-throughput analysis of atomistic protein simulations. *Journal of Bioinformatics and Computational Biology* 10(04) (2012)
6. Benson, N.C., Daggett, V.: A comparison of multiscale methods for the analysis of molecular dynamics simulations. *The Journal of Physical Chemistry B* 116(29), 8722–8731 (2012)
7. Benson, N.C., Daggett, V.: Wavelet analysis of protein motion. *International Journal of Wavelets, Multiresolution and Information Processing* 10(04) (2012)
8. Bowen, J.P., Allinger, N.L.: Molecular mechanics: The art and science of parameterization. *Reviews in Computational Chemistry*, 81–97 (1991)
9. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. *ACM Computing Surveys (CSUR)* 41(3), 15 (2009)
10. Coutsias, E.A., Seok, C., Dill, K.A.: Using quaternions to calculate rmsd. *Journal of Computational Chemistry* 25(15), 1849–1857 (2004)
11. Flocco, M.M., Mowbray, S.L.: C $^{\alpha}$ -based torsion angles: A simple tool to analyze protein conformational changes. *Protein Science* 4(10), 2118–2122 (1995)
12. Guralnik, V., Srivastava, J.: Event detection from time series data. In: *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 33–42. ACM (1999)

13. Kabsch, W.: A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* 32(5), 922–923 (1976)
14. Karplus, M., McCammon, J.A.: Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology* 9(9), 646–652 (2002)
15. Kavraki, L.: Molecular distance measures [connexions web site] (June 2007), <http://cnx.org/content/m11608/1.23/>
16. Kullback, S.: The Kullback-Leibler distance. *The American Statistician* (1987)
17. Ramanathan, A., Agarwal, P.K., Kurnikova, M., Langmead, C.J.: An online approach for mining collective behaviors from molecular dynamics simulations. *Journal of Computational Biology* 17(3), 309–324 (2010)
18. Ramanathan, A., Yoo, J.O., Langmead, C.J.: On-the-fly identification of conformational substates from molecular dynamics simulations. *Journal of Chemical Theory and Computation* (2011)
19. Rapaport, D.C.: *The art of molecular dynamics simulation*. Cambridge Univ. Pr. (2004)
20. Shewhart, W.A.: *Economic control of quality of manufactured product*, vol. 509. American Society for Quality (1980)
21. Smith, A., Datta, S.P., Smith, G.H., Campbell, P.N., Bentley, R., McKenzie, H.A., et al.: *Oxford dictionary of biochemistry and molecular biology*. Oxford University Press (OUP) (2000)
22. Teodoro, M.L., Phillips Jr., G.N., Kavraki, L.E.: Understanding protein flexibility through dimensionality reduction. *Journal of Computational Biology* 10(3-4), 617–634 (2003)
23. Wriggers, W., Stafford, K.A., Shan, Y., Piana, S., Maragakis, P., Lindorff-Larsen, K., Miller, P.J., Gullingsrud, J., Rendleman, C.A., Eastwood, M.P., et al.: Automated event detection and activity monitoring in long molecular dynamics simulations. *Journal of Chemical Theory and Computation* 5(10), 2595–2605 (2009)
24. Zliobaite, I.: *Learning under concept drift: an overview*. Technical report, Technical report, Faculty of Mathematics and Informatics, Vilnius University: Vilnius, Lithuania (2009)