Constructing Comprehensive Behaviors: A Simulation Study

Thomas C. Henderson and Xinwei Xue
School of Computing
University of Utah
Salt Lake City, UT, USA
Email: tch@cs.utah.edu

***Abstract*— A general framework is proposed for behavior modeling, analysis and synthesis. This allows a better understanding and evaluation of the nature and role of behavior models in the following disciplines: ethology, robot behavior specification, animated character behavior specification, and automatic behavior analysis. In addition, it is argued that the following aspects of a system's behavior usefully characterize behavior models: (1) physical, (2) physiological, (3) contextual, and (4) conceptual. We demonstrate how basic behavior units can be extracted from time sequence data of a synthetic two-state problem. This is done by detecting Basic Behavior Units using the affinity graph method, and then determining higher level behavior model parameters based on these.**

I. Introduction

Our goal in this study is to better understand the nature of behavior and behavior models in their many forms. This ranges from the behavior of physical systems according to the laws of physics and chemistry up to the goal-directed behavior of animals, people and autonomous agents. What does it mean for a system to *behave*? Does it mean, as in normal human discourse, to follow the rules? What rules? Are these encoded in the laws of nature, genetics, etc., or merely convention? How do these various aspects interact? Different disciplines take different approaches to formulating and answering these questions. We propose a conceptual framework for this discussion, a set of dimensions of interest in characterizing behavior, and present a method for the discovery of Basic Behavior Units (BBUs), and higher-level behavior model parameters based on them.

The study of the behavior of natural systems is the basic undertaking of science, and the general goal is to produce a description that not only explains what happens, but that can be used to predict future events. For example, a description of the change in height of an object dropped from the top of a building might be derived from Newton's laws and given as a function of height versus time. The behavior in this case is the change in position, and the resulting equation models this behavior. Such a model can be put to a variety of uses; e.g.:

- **explain behavior**: determine time or velocity of impact,
- **predict behavior**: given a desired time of impact, determine the necessary initial height and velocity, or
- **characterize behavior**: given a trajectory, determine if the object obeys the model.

The variables of such models are usually physical quantities that can be measured by well-defined instruments. The result of such measurements is called *raw experimental data*.

A similar approach may be taken in the study of living organisms as in ethology [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. Here the situation is more complicated because behavior is mediated not only by physical laws, but also by physiological conditions, internal drives and environmental context. Also complicating the issue is the interplay between success and survival at the individual and species levels.

In addition, the description of animal behavior may be couched in special variables defined by the investigator and discerned through the psychological processes of the human observer. For example, a gorilla may be watched to determine how often it displays *affection* for its young; a videotape of this would be raw experimental data, but a human produced log of *affection events* based on the video will be termed *annotated behavior* and serves as an explanation of the observed data. Such an explanation is mediated by and couched in terms of the conceptual model.

In order to produce a life-like animation, it is necessary to produce both physically and psychologically correct behavior [11]. Models for animated characters

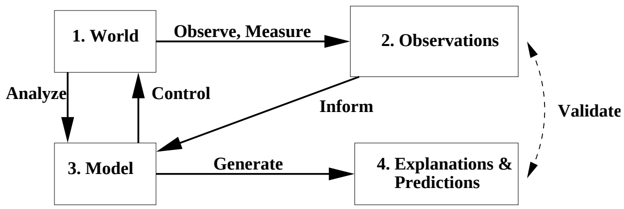

Fig. 1. General Framework for Scientific Explanations

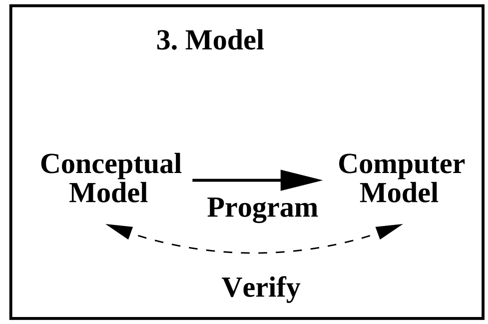

Fig. 2. Computer Models Must Undergo Verification with Conceptual Model.

require a *body* component and a *mind* component. The latter addresses goals, drives, beliefs, etc. A motion sequence generated by such a model will be called a *generated behavior* and is a predicted sequence of events.

The mobile robot research community also produces generated behaviors [12]. However, unlike the animation characters which exist only in an electronic world, physical robots exist in the real world. Thus, these behaviors also include a control aspect in terms of the robot acting in the physical world. (While it is true that an animated character interacts with its virtual world, this again involves generated behaviors, whereas the mobile robot gets physical feedback.)

Finally, the area which most interests us is automatic behavior analysis. Here the goal is to combine raw experimental data (usually video) with a behavior model and produce what we term *interpreted behavior*. This corresponds to annotated behavior except that one is produced by humans and the other by computation. Interpreted behavior thus also serves as an explanation of the observations in terms of the model.

To better understand the various manifestations of behavior and the emphases of different disciplines, we propose the general framework for scientific investigation shown in Figure 1. The world (Box 1) signifies the object of study which may be, for example, gravitational force, or the foraging behavior of army ants. Typically, direct access to the world is not possible (i.e., through Plato's ideals, or Kantian categories), and the world must be understood through observation and measurement (Box 2). Such observations may arise through human perception or the use of measuring instruments.

A model (Box 3) of the object of interest is developed based on measurements and observations of the object. Modeling and observation are highly coupled in that the observations provide desiderata for model creation, while the model itself informs the experimental framework for data acquisition. The model serves two major purposes; first, it should explain the observations; second, it should predict new phenomena. These explanations and predictions (Box 4) can be compared to the observations in order to **validate** the model. Finally, the model can provide guidelines to control the object of study; this can either be to define or improve observation conditions, or can be done with the goal of achieving a certain predicted result.

Another level of detail is required to distinguish computer models from other formal frameworks; this is shown in Figure 2. The conceptual model is converted to a computer model by programming an implementation. To ensure the equivalence of the two models requires **verification**. This includes, understanding and eliminating algorithmic errors, numerical errors, coding errors, etc.

II. Behavior Modeling and Analysis

The automatic analysis of behavior requires:

- **model building**: theory and/or observations are used to determine the component models (physical, physiological, conceptual).
- **model exploitation**: each model is combined with the appropriate observations to explain (interpret) the observables.

Model exploitation follows the framework given in Figure 1 and provides explanations (interpretations) of the observables in terms of the behavioral units of the model. However, model building is not so easily accomplished. As pointed out by Colgan[13], a *model;* consists of two components:

1) a description of a mathematical (or other) system, and
2) a map that links the variables of (1) with those of the object under study.

It is often difficult to determine appropriate variables and mappings. One major issue in model building is that the observables may only be loosely connected to the deeper behavior-driving phenomena. This leads to the creation of models defined in terms of superficial qualities which only indirectly allow understanding of the behavior generator.

Another important aspect of a modeling effort is the impact of error at various stages of the process. It is important to quantify error and how it propagates through the model into the interpretation. Of course, standard error analysis can always be applied (e.g., see [14]), but it is also necessary to understand discontinuous error. For example, if the desired explanation is a label like *grooming*, then it is either correct or not, and it may be difficult to relate smoothly to the observation variables derived from images. However, it is precisely this type of sensitivity analysis which is needed and which should be supported by the modeling technique. The simulation community has developed a Verification and Validation approach to the development of complex, multi-physics, multi-scale codes [15], and we propose to incorporate some of their ideas here.

We propose to construct a general behavior model whose variables and their relations describe the following aspects of behavior:

1) physical: any known physical laws of mathematical models (physics).
2) physiological: any rules relating body regulation, etc. (chemistry and physiology).
3) contextual: space or time circumstances which influence behavior (reactive).
4) conceptual: mental states or properties (AI).

Variables may be discrete or continuous, temporal, spatial, internal or external, or abstract. What differentiates models is the nature of the variables and their relationships to each other and the modeled system.

Once a model has been determined, it is useful to evaluate it. We believe that the following 2 aspects are the most important:

- How well does it describe the 4 model levels given above (i.e., physical, physiological, contextual, and conceptual)?
- How well does it enable *arrow activity* in Figures 1 and 2? Examples of this include: (1) How well are observations used to inform the model building? (2) What is the complexity of generating explanations or predictions? (3) How easily can validation be performed and at what accuracy?[1]

We now turn our attention to the various forms of conceptual models for behavior. Our approach to

[1]The authors would like to thank Ann Torrence for this observation.

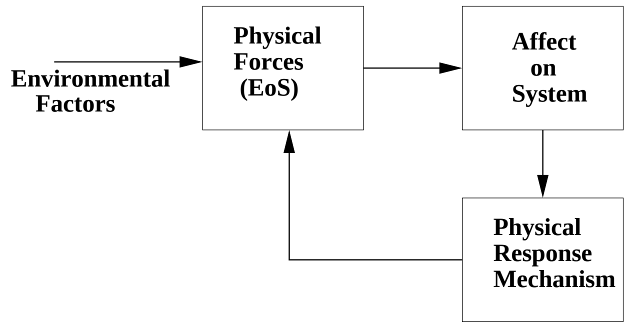

Fig. 3. Physical System Behavior Model.

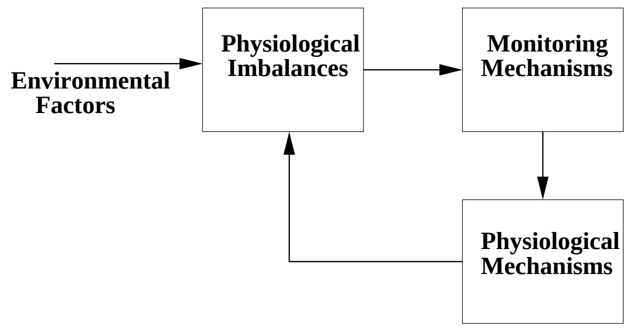

Fig. 4. Physiological System Behavior Model.

modeling follows that of McFarland [16], [17]. The basic model for the *physical system* behavior is shown in Figure 3. The model is divided into two major parts: (1) the Equations of State (EoS) which describe all forces of interest at work in the system, and (2) the specific characteristics of the particular object under study. For example, (1) will usually elaborate $F = ma$ while (2) specifies mass, initial position velocity, etc., as well as any other local constraints (e.g., gravitational constant, existence of floors, walls, etc.). (This area has been well studied and is not addressed here.)

Since we are interested in biological systems, the next level of behavior model describes the *physiological system*. Figure 4 shows the basic scheme for this, and it is much like the physical system. Although at some level of description this is perhaps just a physical system, the scale at which events of interest occur is too far removed from physical forces to be modeled that way. For example, we may want to model hunger as a drive, and this approach allows an appropriate conceptualization of hunger.

Finally, the full model for autonomous agents is given in Figure 5. The physical and physiological systems are integral components of this model. The behavioral mechanisms (i.e., the action generating processes) give the possible responses of the system. Such actions have consequences both in terms of the

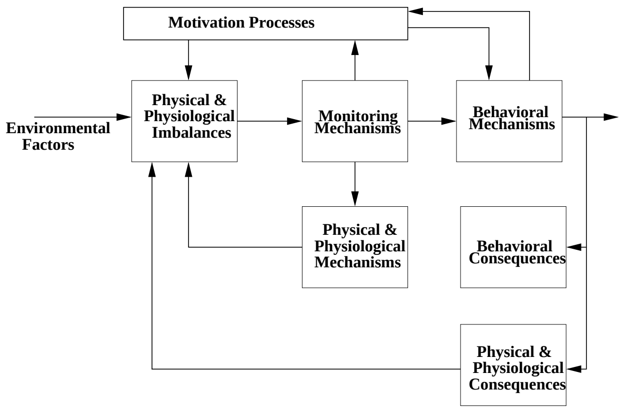

Fig. 5. Motivated System Behavior Model.

behavioral state of the agent, as well as in terms of the physical and physiological state. For example, if the selected action is *eat*, then there are required physical motions, and there are physiological consequences such as decrease in hunger and increase in thirst.

The *motivational processes* are those that play a role in creating goals, shifting attention, influencing drives, etc. and which are not strictly physical or physiological. Such processes may interact intimately with lower level processes; for example, the agent may choose to ignore pain in order to obtain food.

Note that although learning is a major aspect of behavior, it is not addressed here. We hope to take this up in a future study.

III. Basic Behavior Unit Discovery

A. Approach

A behavior model is built in terms of basic behavior units (BBUs). An appropriate BBU set must be found for the particular modeling approach. For example, suppose that we wish to model a mouse in a cage. Then, a set of BBUs of interest might include: resting, eating, drinking, grooming, and exploring. If we adopt the state-space approach, then the observable variables we will use are: position ($p(t)$), speed ($s(t)$), and acceleration ($a(t)$).

It is possible to make general functional characterizations of the BBUs in terms of the temporal variation of these variables. For example:

- resting: $p(t) =$ ground level; $s(t) = 0$; $a(t) = 0$
- eating: $p(t) =$ raised body; $s(t) = 0$; $a(t) = 0$
- drinking: $p(t) =$ raised body; $s(t) = 0$; $a(t) = 0$
- grooming: $p(t) =$ any; $s(t) = sin(t)$; $a(t) = square(t)$
- exploring: $p(t) =$ any; $s(t)$ varies randomly; $a(t)$ varies randomly

However, this is difficult since it involves high level notions about motion (random, sine, square wave, etc.), and in fact, should consider the motions of the limbs and head separately. Another approach is to obtain video data of the BBUs of interest, and then calculate time sequences of position (e.g., of the center of mass), speed, and acceleration, and determine whether these allow discrimination of the distinct BBUs. We follow this approach.

In related work, Jenkins [18] employs a spatiotemporal nonlinear dimension reduction technique (PCA-based) to derive action and behavior primitives from motion capture data, for modularizing humanoid robot control. First, spatiotemporal neighborhoods are built, then a matrix D of all pairs shortest distance paths is computed, and finally PCA is performed on matrix D. Barbic et al. [19] propose three PCA-based approaches. The approach we use, the affinity graph method, has mostly been applied in image segmentation, as summarized in [20]. Recently, this method has been applied to event detection in video [21], [22]. Others have used the concept of similarity matrix for classification (e.g., gait recognition [23] and action recognition [24]).

Different affinity measures have been proposed to construct the affinity matrix. In image segmentation, distance, intensity, color, texture and motion have been used [25]. In video-based event detection, a statistical distance measure between video sequences is proposed based on spatiotemporal intensity gradients at multiple temporal scales. [22] uses a mixture of object-based and frame-based features, which consist of histograms of aspect ratio, slant, orientation, speed, color, size, etc., as generated by a video tracker.

Most closely related to our approach are [21] and [22]. [21] constructs affinity matrix from temporal subsequences using a single feature, while [22] constructs the affinity matrix for each frame based on multiple weighted features.

We are particularly interested in discovering animal behaviors from video sequences, and use the affinity graph method to segment temporal sequences into BBUs. This is done by choosing an element for consideration. For us, this element is a temporal subsequence of length T. Next a matrix is constructed in which each $(i; j)$ entry gives an affinity (or similarity) measure of the i^{th} and j^{th} elements (we use the element-wise distance between the subsequences). The eigenvalues and eigenvectors of the matrix are found, and the eigenvalues give evidence of the strength of a cluster of similar elements. As described in [25], if we maximize the objective function $w_n^T A w_n$ with affinity matrix A and weight vector w_n linking elements to the n_{th} cluster, and requiring $w_n^T w_n = 1$, then the Lagrangian is:

$$w_n^T \mathcal{A} w_n + \lambda(w_n^T w_n - 1)$$

which leads to solving $\mathcal{A} w_n = \lambda w_n$. Therefore, w_n is an eigenvector of $\mathcal{A}$. Thus, the eigenvectors of the affinity matrix determine which elements are in which cluster. We use this to extract basic behavior units in terms of their position, velocity, etc. of various state variables of interest.

Given a segmentation of the behavior sequence into BBUs, the next goal is to determine higher-level model parameters and structure. For example, we might assume a Hidden Markov Model approach, or some parameterized characteristic function, and use the temporal relations between the BBUs to form a concrete model.

B. Simulation of A Simple Two-State Example

Data Synthesis. For purposes of demonstration, we model some very simple behaviors of a mouse in a cage. Here we assume that the physiological, contextual and conceptual models may be expressed as probabilistic functions of some S-curve form (sigmoidal, hyperbolic tangent, etc.). As the Basic Behavior Units, suppose the mouse can either rest (BBU_{rest}) or wander (BBU_{wander}). Furthermore, suppose that the transition between these two behaviors is characterized by two functions, $F_{rest \rightarrow wander}(t)$ and $F_{wander \rightarrow rest}(t)$:

$$F_{rest \rightarrow wander}(t) = 1/(1 + e^{K_{rest} - t})$$

$$F_{wander \rightarrow rest}(t) = 1/(1 + e^{K_{wander} - t})$$

where K_{rest} is a parameter specifying the length of rest periods, and the function gives the likelihood as a function of time that the mouse will wake up and start to wander. K_{wander} is a parameter specifying the distance wandered, and gives the likelihood that after moving a distance d the mouse will stop wandering and begin to rest. Figure 6 shows the transition likelihood for the sigmoid models used here to synthesize data sequences. Behavior sequences are synthesized over 20,000 time steps using fixed values of K_{rest} = 40 and K_{wander} = 40, and the observables are: $x, \dot{x}, y, \dot{y}, a, \dot{a}$, where a is the mouse heading angle.

Data Analysis. First, the basic behavior units (rest, explore) are determined using the affinity graph method. Figure 7 shows the segmentation of part of the data sequence (the actual behavior sequence is shown as positive values and the segmented as negative to enhance the visual effect). A critical parameter

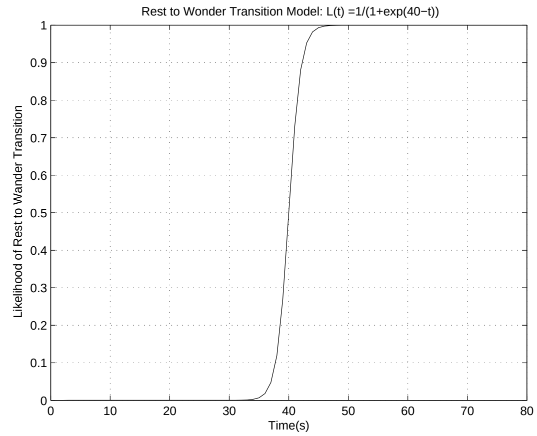

Fig. 6. Likelihood of transition from rest to wander behavior function (sigmoid)

in this temporal sequence analysis is the subsample time period, T (set to 3 here). As can be seen in the figure, there is a little error at the onset of each behavior segment. The error in segmentation (i.e., number of time steps incorrectly labeled) is about 3%. We are going to develop a statistical algorithm that can choose the optimal T parameter.

The time spent in individual behavior units is used to develop a statistical model of the transition probability. These probabilities may be used to form different types of models (HMM, etc.); here we recover the same form as was used to generate the data in order to allow a straightforward comparison of the results. The best parameters for $F_{rest \rightarrow wander}(t)$, and $F_{wander \rightarrow rest}(t)$ are then determined. Let C_{rest} be the total number of resting BBUs and C_{wander} be the total number of wandering BBUs.

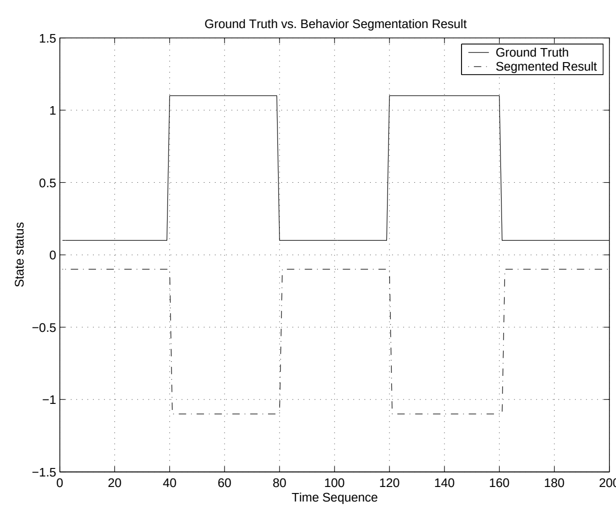

Fig. 7. Comparison of segmented result with ground truth

Transition likelihoods are calculated using the length of time spent in each BBU; let the length of the i^{th} BBU be $|BBU_{state,i}|$; then:

$$L_{rest\to wander}(t) = \frac{|\{|BBU_{rest,i}| < t\}|}{C_{rest}}$$

$$L_{wander\to rest}(t) = \frac{|\{|BBU_{wander,i}| < t\}|}{C_{wander}}$$

Figure 8 shows the histogram of the times spent at rest, and Figure 9 shows the cumulative likelihood of transition curve derived from the histogram (i.e., its integral). $\hat{K}_{rest}$, the estimated value of K_{rest}, is 37.5 (versus 40). Next consider the role of physiological,

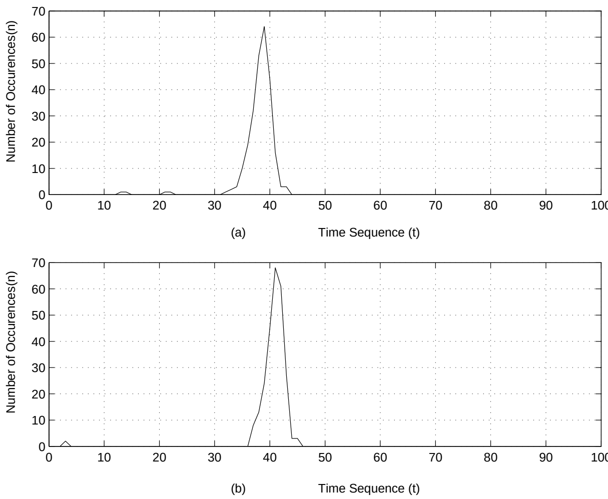

Fig. 8. Histogram of Action Transition Intervals for (a) Rest⇒Explore Transition and (b) Explore⇒Rest Transition

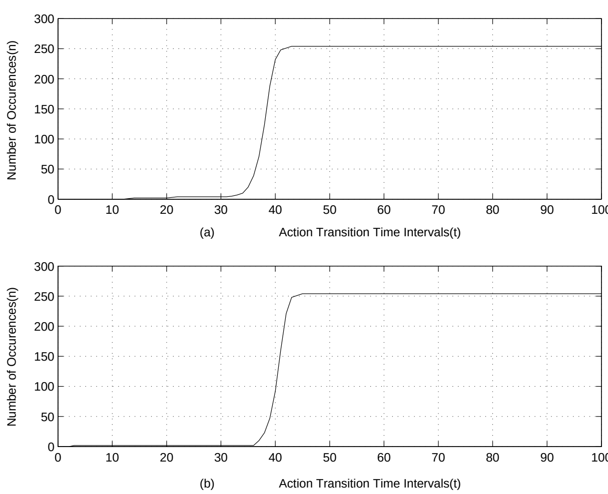

Fig. 9. Cumulative Histogram of Action Transition Intervals for (a) Rest⇒Explore Transition and (b) Explore⇒Rest Transition

contextual or conceptual variables in determining

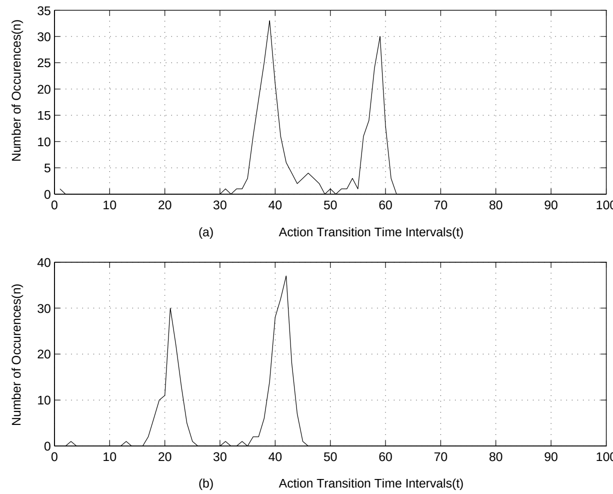

Fig. 10. Histogram of Action Transition Intervals with Light Context, for (a) Rest⇒Explore Transition and (b) Explore⇒Rest Transition

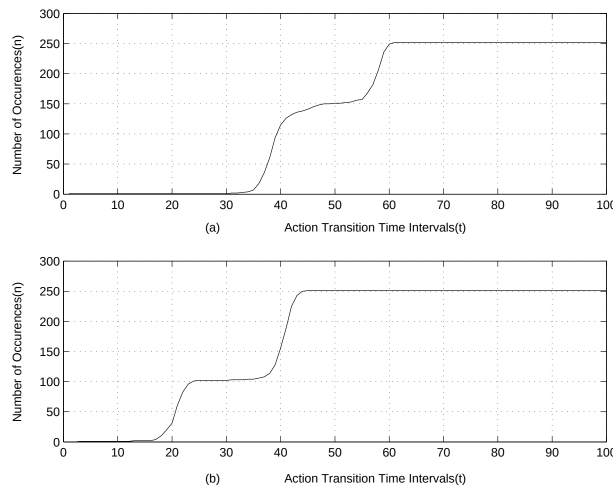

Fig. 11. Cumulative Histogram of Action Transition Intervals with Light Context, for (a) Rest⇒Explore Transition and (b) Explore⇒Rest Transition

behavior. Our premise is that these variables change the parameter or form of the behavior likelihood functions. For simplicity of the demonstration, we assume that only the function parameter changes with the change in physiology, context or conceptual frame of mind. For example, suppose that the mouse tends to rest for longer periods and wander for shorter periods when it is dark; then the resting transition likelihood function shifts to the right, and the wander function to the left. If a behavior sequence is available which includes periods of dark and light, then this is readily determined by the appearance of multiple peaks in the transition time histogram (see Figure 10).

Functions with the appropriate respective parameters for light and dark can then be found. Figure 11 shows this with the shifted versions of the transition likelihood functions. It is also possible to determine the causal role of light if the observed data includes some measure of the phenomenon (e.g., light intensity as a function of time).

For physiological and conceptual variables, there will be no corresponding observable data. However, it is still possible to detect multiple peaks in the behavior time histogram and infer hidden variables.

IV. Discussion and Conclusions

We propose a framework for the study of behavior modeling which includes physical, physiological, contextual and conceptual levels. The affinity graph method is proposed for the segmentation of BBUs based on physical observations. BBUs are required to build higher-level behavior models. We show how BBUs are used to determine parameters at the physiological, contextual, and conceptual levels.

We are currently investigating the application of the method to model the activities of a lab mouse in a cage (Figure 12 shows a mouse in the lower left sitting quietly in the cage). Its activities are

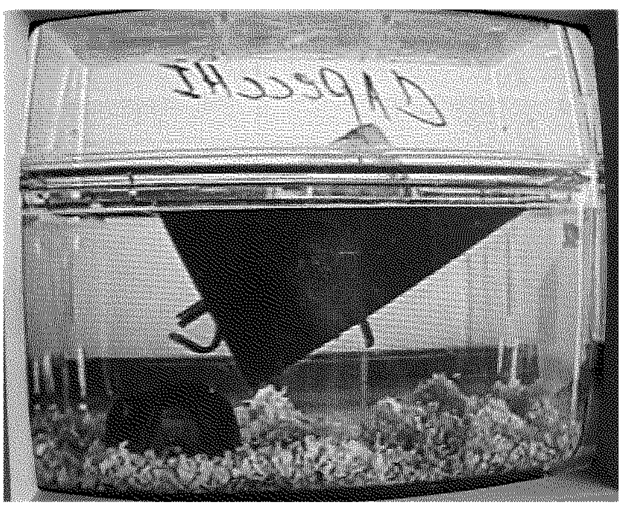

Fig. 12. Mouse in Cage.

observed and recorded manually (eventually we hope to automatically extract this from video sequences). Here we show our preliminary results of finding basic behaviors using the affinity graph. The mouse's behaviors include: exploring, reaching up, grooming, or staying still. The observed data consists of a vector, including position (x,y), and two angles of the head of the mouse relative to its rear body: one to distinguish whether it is standing up or staying on ground, the other to indicate whether it is grooming with its body twisted. Here is a sample of observed parameter sequence:

x	y	theta1	theta2	time	lasts	action
10	5	0	0	0	90	still
8	4	0	0	90	5	move
8	4	90	0	95	57	reach up
9	6	0	0	152	10	down
10	4	0	90	162	30	groom

The observed data consists of parameters recorded in the same style as above. In the analysis, only the first four parameters are used. The data is interpolated to one second samples. We use the affinity graph method to determine which elementary sequences cluster into basic behavior units. There are 11 major eigenvalues found – just as there are about this many distinct behaviors for the mouse (of course, the behaviors can be divided into different length time sequences to get different numbers of behaviors). The BBU detection seems adequate, and in future work aim to achieve comprehensive behavior model construction for such scenarios.

Acknowlegments

This research is funded by the University of Utah Center for the Simulation of Accidental Fires and Explosions (C-SAFE), funded by the Department of Energy, Lawrence Livermore National Laboratory, under subcontract B524196.

References

[1] B. A. Hazlett, *Quantitative Methods in the Study of Animal Behavior*. New York: Academic Press, second ed., 1977.
[2] M. Mangel and C. W. Clark, *Dynamic Modeling in Behavioral Ecology*. Princeton, NJ: Princeton University Press, 1988.
[3] F. M. Toates, *Animal Behavior: A Systems Approach*. New York: John Wiley & Sons, 1980.
[4] I. Eibl-Eibesfeldt, *Ethology: The Biology of Behavior*. Holt, Rinehart and Winston, second ed., 1975.
[5] C. W. Clark and M. Mangel, *Dynamic State Variable Models in Ecology: Methods and Applications*. New York: Oxford University Press, 2000.
[6] J. Bart, M. A. Fligner, and W. I. Notz, *Sampling and Statistical Methods for Behavioral Ecologists*. New York: Cambridge University Press, 1998.
[7] J. M. Gottman and A. K. Roy, *Sequential Analysis: A Guide For Behavioral Researchers*. Cambridge: Cambridge University Press, 1990.
[8] P. Martin and F. P. Bateson, *Measuring Behavior: An Introductory Guide*. Cambridge,UK: Cambridge University Press, second ed., 1993.
[9] R. Bakeman and J. M. Gottman, *Observing Interaction: An Introduction to Sequential Analysis*. Cambridge, UK: Cambridge University Press, second ed., 1997.
[10] L. A. Dugatkin and H. K. Reeve, *Game Theory and Animal Behavior*. Cambridge, UK: Oxford University Press, 1998.
[11] H. Prendinger, *Life-Like Characters: Tools, Affective Functions and Applications*. New York: Springer, 2003.
[12] R. C. Arkin, *Behavior-Based Robotics*. Cambridge, Massachusetts: The MIT Press, 1998.

[13] P. W. Colgan, *Quantitative Ethology*. New York: John Wiley & Sons, 1978.

[14] J. R. Taylor, *An Introduction to Error Analysis*. Sausalito, CA: University Science Books, 1997.

[15] W. Oberkampf and T. Trucano, "Verification and validation in computational fluid dynamics," Sandia National Laboratory Report SAND2002-059, Sandia Labs, 2002.

[16] D. McFarland and A. Houston, *Quantitative Ethology: The State Space Approach*. Boston,MA: Pitman Advanced Publishing Program, 1981.

[17] O. Holland and D. McFarland, *Artificial Ethology*. New York: Oxford University Press, 2001.

[18] O. C. Jenkins and M. J. Mataric, "Deriving action and behavior primitives from human motion data," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS)*, pp. 2551–2556, 2002.

[19] J. Barbic, A. Safonova, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. S. Polland, "Segmenting motion capture data into distinct behaviors," in *Proceedings of Graphics Interface 2004 (GI'04)*, (Canada), May 2004.

[20] Y. Weiss, "Segmentation using eigenvectors: a unifying view," in *Proc. IEEE International Conference on Computer Vision*, pp. 975–982, 1999.

[21] L. Zelnik-Manor and M. Irani, "Event-based analysis of video," in *IEEE CVPR*, 2001.

[22] F. Porikli and T. Haga, "Event detection by eigenvector decomposition using object and frame features," in *Workshop on Event Mining, IEEE ICCV*, 2004.

[23] C. BenAbdelkader, R. G. Cutler, and L. S. Davis, "Gait recognition using image self-similarity," *EURASIP Journal on Applied Signal Processing*, vol. 4, pp. 572–585, 2004.

[24] A. A. Efros, A. C. Berg, G. Mori, and J. Malik, "Recognizing action at a distance," in *IEEE International Conference on Computer Vision*, (Nice, France), pp. 726–733, 2003.

[25] D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall, 2003.