Social scientists study the dynamics of self-organizing dyadic processes.

For humans, marital relationships have a direct effect on health. Good relationship protects where as a bad one has negative effects on health. It also causes psychological problems with children.

From the mid 1970s investigators had hoped to establish a set of critical macro-level behavior patterns, derived from micro-social processes, that predict marital quality. This study shows consistent, albeit slight, structural differences at each level of aggregation. Their study was on analyzing, illustrating, and classifying one or two modalities of behavior.

Over the last few decades most of the significant insights into couple dynamics arose not from better data or measurement methods but improved data analytic strategies. The area moved steadily from summary statistics describing individual behavior during an interaction (e.g., mean number of positive statements) to probability estimates of shifting dyadic structures (e.g., sequential analysis) to state transition models

algorithms that generate the evolving contingency structures embedded in sequential data are not well understood. In turn, this has severely limited attempts to model dyadic interaction

To construct realistic and informative generative models of dyadic interactions three analytic problems need addressing:

(1) how to articulate the state space of the dyad

(2) how to generate a tenable state transition matrix

(3) how to incorporate duration expectancy into states and transitions.

In practice, Affective Computing is a scientific chimera that integrates existing literature on verbal and non-verbal social behavior, such as that cited above, incorporates ongoing laboratory studies on affect (e.g., computer-user interfaces) and builds computational models using methods drawn from the machine learning field

t’s this latter ability that is lacking in the existing marital interaction literature—reinvigorating this area requires moving from the language of social science and into the realm of computer science: specifically, to advance our understanding of intimate dyadic affective and behavioral structures we need generative models of sequential latent processes, i.e., profiles of sequential movement across latent states with estimated probabilistic structures

With the advent of faster computers, and theoretical advances in Bayesian methods, emphasis shifted toward discovering methods that capture temporal clustering and feature extraction in sequential data. At the forefront of this area is the family of nonparametric Bayesian techniques. For example, Hierarchical Dirichlet Process methods permit investigators to partition data into topics and sequence sensitive feature states

Initially, HMMs were used as tools for voice recognition and only later adopted for myriad recognition tasks, ranging from social dynamics and biological sequences to economic and hydrological time-series. Recently HMM methodologies began incorporating the capacity of nonparametric Bayesian approaches to define prior distributions on transition matrices over countably infinite state spaces; adopting this technique allows a greater range of use with real, somewhat messy, data

Griffin collected 30 couples data to develop 10 state 4 symbol HMM which had classification of 91%.

Three aspects of this research were noteworthy.

1. The classification rate of 91% is below expectation in a well developed HMM, but this value is acceptable, especially in this social science area, given the small sample size and low dimensional vector used to create the data string.
2. Although no substantive conclusions were forwarded, results demonstrated that couple interactions were patterned and that accurate machine learning and classification can occur without supervision.
3. When these analyses occurred, Hidden semi-Markov Models, although well articulated in theory and few examples were in the literature, practical implementation was years away

Increased investigations of real-world complex datasets, the standard HMM has been transformed over the last decade—computer scientists created numerous sequential analytic techniques that are sensitive to the nuances of evolving latent structures (e.g., infinite HMMs) akin to the type seen in micro-social dynamics

The work in the article incorporates the methodology developed by Johnson and Willsky; our goal of extracting patterns from sequential data is conceptually and quantitatively similar to their search for structure in real and synthetic data. Whereas they used multiple time-series of household appliance data, we inserted husband and wife sequence data, and we have added the additional dimension of relationship satisfaction as a classification problem. Additionally, our simulations are written in Python, as is their publicly available code, thus it was modified to fit our research questions, output plots, and GUI development.

To use the HDP-HSMM methodology required that we partition the couples into homogeneous sub-groups.

hierarchical clustering techniques to construct the sub-groups

Collected data by posting an ad in newspapers and offered 25$. 19 couples were selected and 11 were selected later.

method of metric retrieval, each effect has a subjective reference that is unique to the rater, within the context of the interaction, given the dyad’s history. For each individual there is only an internal template referencing their affective state; an internal state that is pleasant to one individual may be only neutral to another. Moreover, because it is self-report, it is arguable that such a recall procedure provides a good proxy of the true affect state, and requires less interference than other, outsider perspective data collection procedures.

based on ample evidence, that happily married couples interact differently than those who are dissatisfied

Historically couples were classified as either distressed or non-distressed.

**Means & Standard Deviations** Two additional features were also taken from the States variable; the first is the average States score, values ranged from 2.001 to 7.467 (*μ* = 5.336, *σ* = 1.463); lower scores indicate greater positive affect. Second, we used the States standard deviation; values ranged from 0.654 to 2.282 (*μ* = 1.479, *σ* = 0.445). Inter-feature correlations are shown in **note** that, not surprisingly, entropy and the States standard deviation are highly correlated, others are in the expected direction and range.

DP-HSMM extends the traditional Hidden Markov Model in two fundamental ways:

(1) by incorporating varying state durations, micro-social event dynamics are not constrained to a geometric form—acknowledging that time in state makes a difference in sequential behavior

(2) by allowing a countably infinite number of states the model incorporates dyadic histories with ideographic state spaces—each dyad has a unique number of states that best capture their behavioral propensities.

## **Results**

### **Clusters**

Each feature (*MAT Score*, *States mean*, *States standard deviation*, *Shannon entropy*, *Dynamic Time Warping*) was initially normalized (0,1) using the MinMax method. After normalizing the data, a series of pair-wise euclidean distance estimates were taken (row 1 vs row 2, etc). This normalized distance matrix was then used to construct clusters.

To generate clusters we employed hierarchical clustering techniques, alternating between average linkage and Ward’s method

Similar result was observed.

### **HDP-HSMM Fit**

A summary of the results showed minimal differences between the Hamming and the Dynamic Hamming distances, thus for we only report the Hamming distance. With one exception, all best fit combinations were with 20 States. Using the best fitting model criteria, aside from maximum number of States, no other single parameter (e.g., *α* or *γ* in the transition distribution) predominated, either at the couple level or within Cluster