**Model**

Beta Process HMM [Fox et al. NIPS ’09]

Generates collections of sequential data

Global set of behaviors (features) # of behaviors learned from data

\[ \theta \]

emission parameters

Each sequence uses a sparse subset of behaviors

\[ \{z, \theta \} \]

features

\[ \text{RARELY} \]

creates new features (f)

---

**Data-Driven Birth/Death**

Add or delete unique feature to one sequence

Reversible jump proposal for \( F, \theta \) (marginalized out)

\[ F^{x} \]

death

\[ F^{y} \]

birth

Propose from prior (Fox et al. NIPS 2009)

\[ \theta_k^{x} \sim \frac{1}{2} \phi(\theta) + \frac{1}{2} \phi(x_i : i \in W) \]

Data-driven proposal

\[ \theta_k^{y} \sim \frac{1}{2} \phi(x_i : i \in W) \]

Select random word \( W \) of sequence

\[ \theta_k \]

selects mixture of prior and posterior over \( W \)

\[ \theta_k \]

appears in \( \{x_i \} \)

\[ \theta_k \]

observed data at time \( t \) in seq.

\[ x_i \]

active feature at time \( t \) in seq.

\[ \phi(\cdot) \]

activates good death move acceptance rate

\[ \phi(x_i) \]

efficiently adds new behaviors informed by the data

---

**Toy Experiments**

Existing MCMC inference slow to mix

\[ \text{Requires long time to make big changes} \]

Each update touches small subset of variables

\[ \text{Rarely creates new features} \]

Proposals from vague prior poorly matched to data

**Contributions**

- Split-Merge (SM) move
  - Many more feature assignments at once
- Data-Driven (DD) birth/death move
  - Propose new features consistent with observed data

\[ \theta_k \]

Significantly improves mixing.

Scales to 100+ sequences.

---

**Split-Merge**

Yields exact posterior samples with big changes to feature assignments

via reversible Metropolis-Hastings proposal moves.

\[ \text{For each seq. in } \mathcal{S}, \text{ } n \in S \]

\[ \text{must own } k_{i}, \text{ black sampled given } F \]

Features available in \( k_{i} \) can appear anywhere in \( z \)

\[ \text{with move detailed in Figure 3} \]

\[ \text{acceptance ratio set to } \text{arbitrary} \]

\[ \text{settings in Table 1} \]

\[ \text{MERGE proposal} \]

\[ \text{SPLIT proposal} \]

\[ [\text{MH acceptance ratio for }] \]

\[ \text{split proposal} \]

\[ \text{MERGE proposal} \]

\[ [\text{MH acceptance ratio for }] \]

\[ \text{merging proposal} \]

\[ \text{Feature selection} \]

---

**Problem**

Existing MCMC inference slow to mix

\[ \text{Requires long time to make big changes} \]

Each update touches small subset of variables

\[ \text{Rarely creates new features} \]

Proposals from vague prior poorly matched to data

**Contributions**

- Split-Merge (SM) move
  - Many more feature assignments at once
- Data-Driven (DD) birth/death move
  - Propose new features consistent with observed data

**Motion Capture**

6 subjects performing 12 exercises (7 shared, 5 unique).

Data: 12 joint angle sensors. Model: 1st-order auto-regressive.

Goal: Compare BHM's segmentation to human annotation.

Previous work stuck with split behaviors even with clever initialization.

New MCMC finds better segmentation starting with just one feature.

Example: discovered behaviors on larger 124 sequence dataset

<table>
<thead>
<tr>
<th>SM and DD moves find essential features</th>
<th>Previous methods stuck in bad local optima</th>
</tr>
</thead>
</table>

**Kitchen Video**

126 subjects prepare 5 recipes in common kitchen. Model: Multinomial.

Data: Bag of visual words (motion + appearance), at 1 second intervals. Unsupervised.

Example: key frames:

- Open Fridge
- Light Switch
- Stir Blender

---