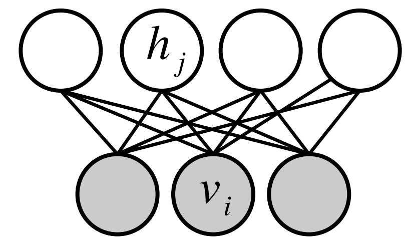
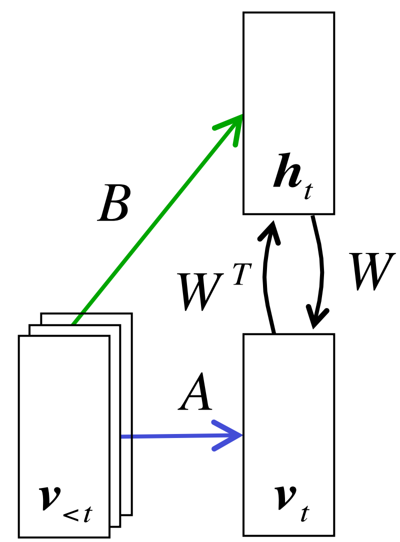
Factored Conditional Restricted Boltzmann Machines for Modeling Motion Style

Graham Taylor and Geoffrey Hinton
University of Toronto, Canada

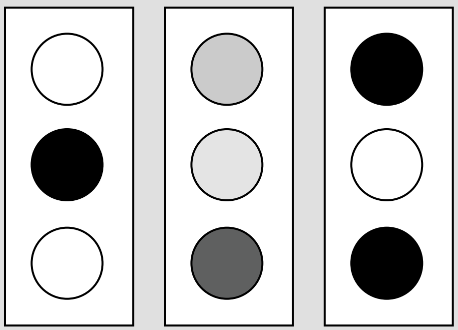
Conditional Restricted Boltzmann Machines

- Start with an RBM (binary-binary or real-binary)
- Add two types of directed connections
- Does not change inference and learning
- Autoregressive weights model short-term, linear structure



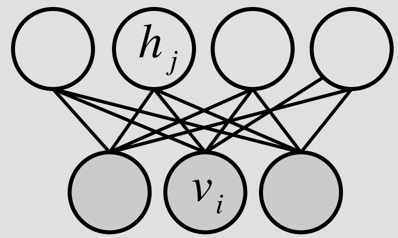


Distributed



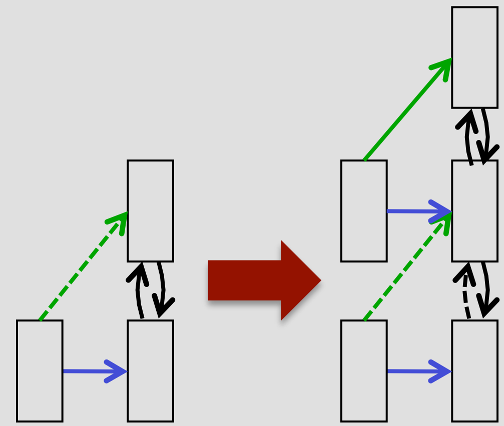
- Capable of representing data that is a product of multiple underlying influences

Undirected



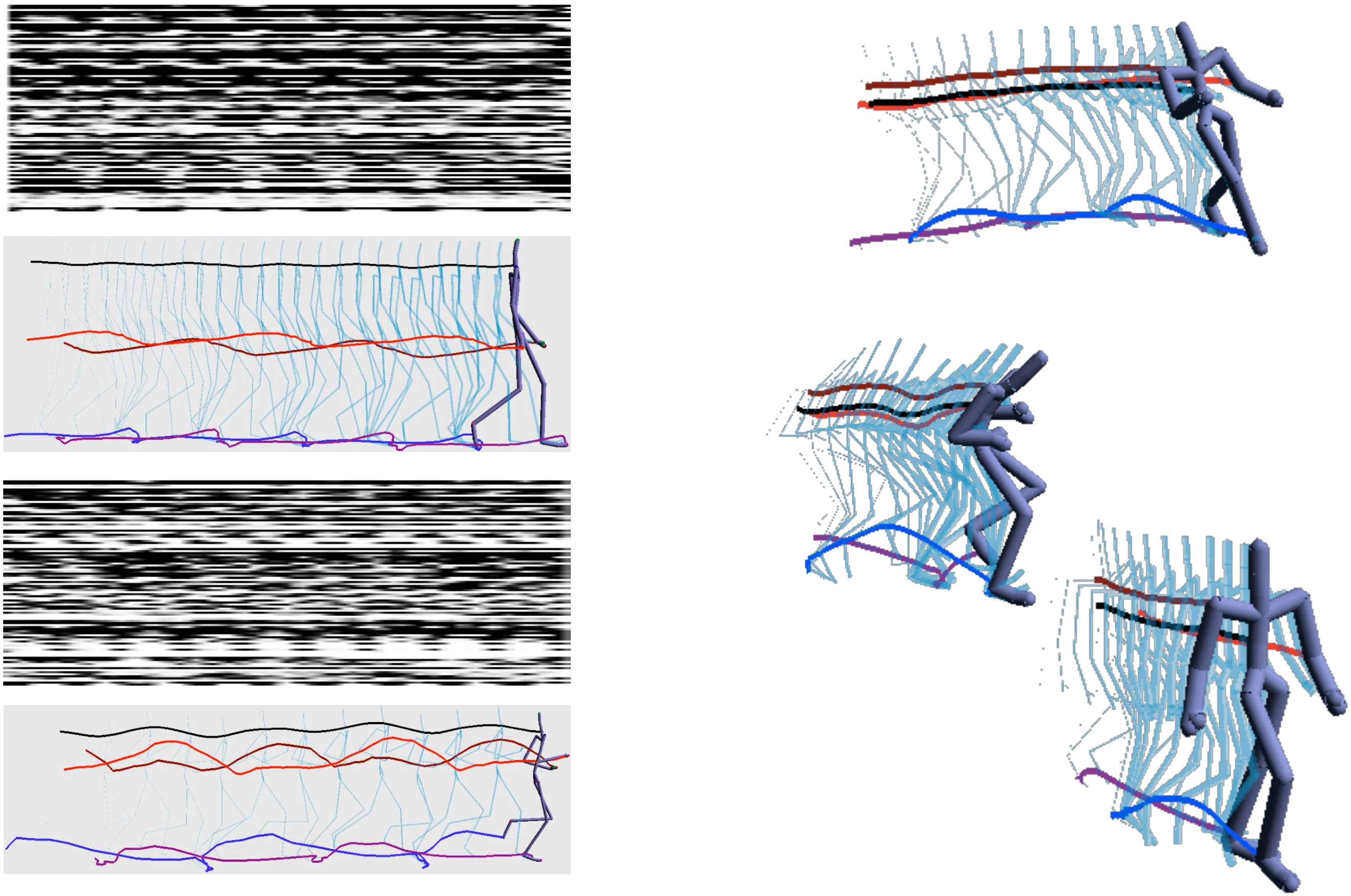
- Using an RBM makes exact inference easy

Composable

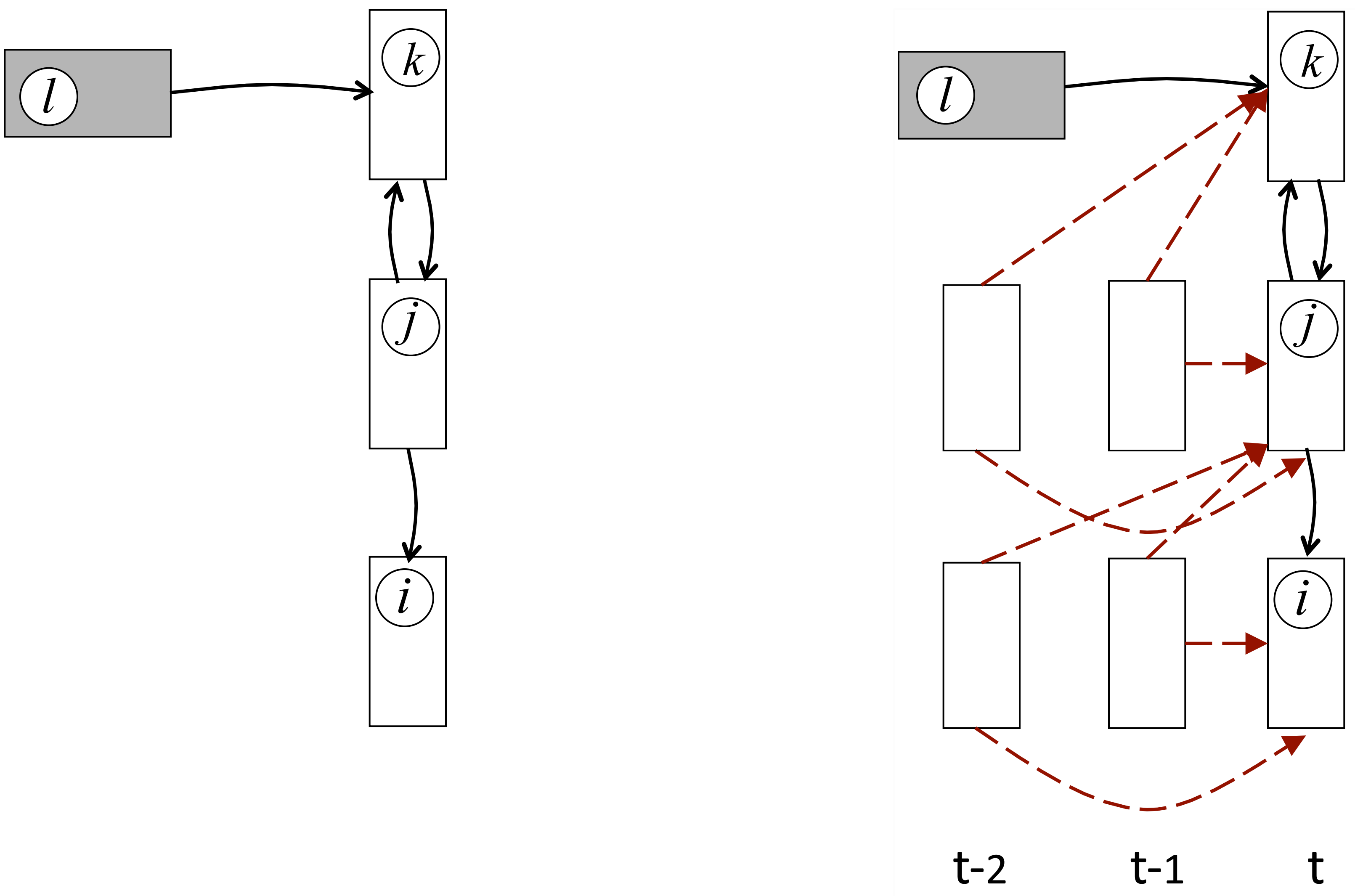


- Train greedily, layer-by-layer

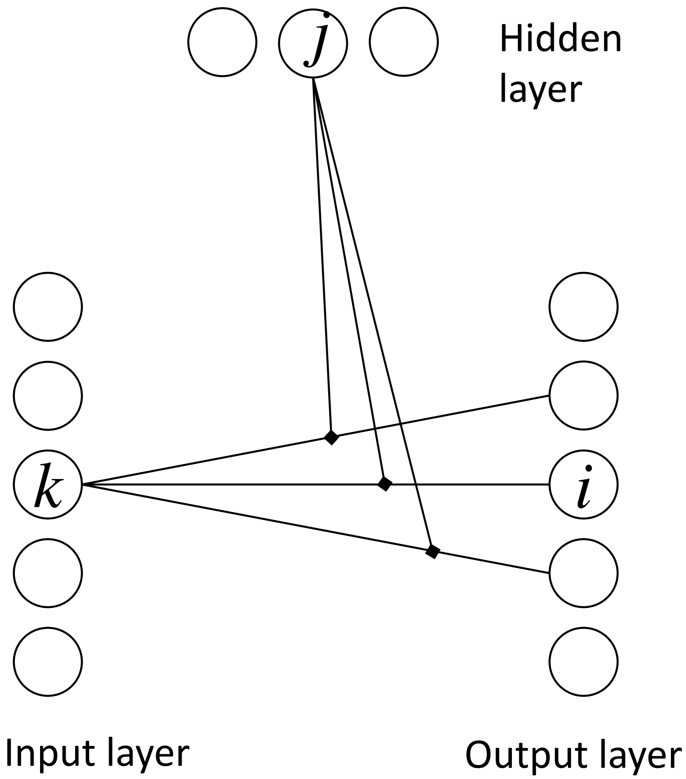
Modeling multiple styles of motion

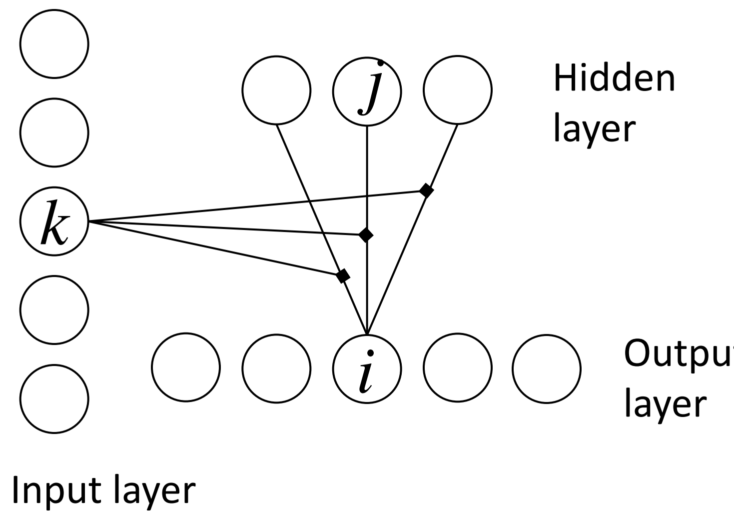


Learning style



Higher-order interactions

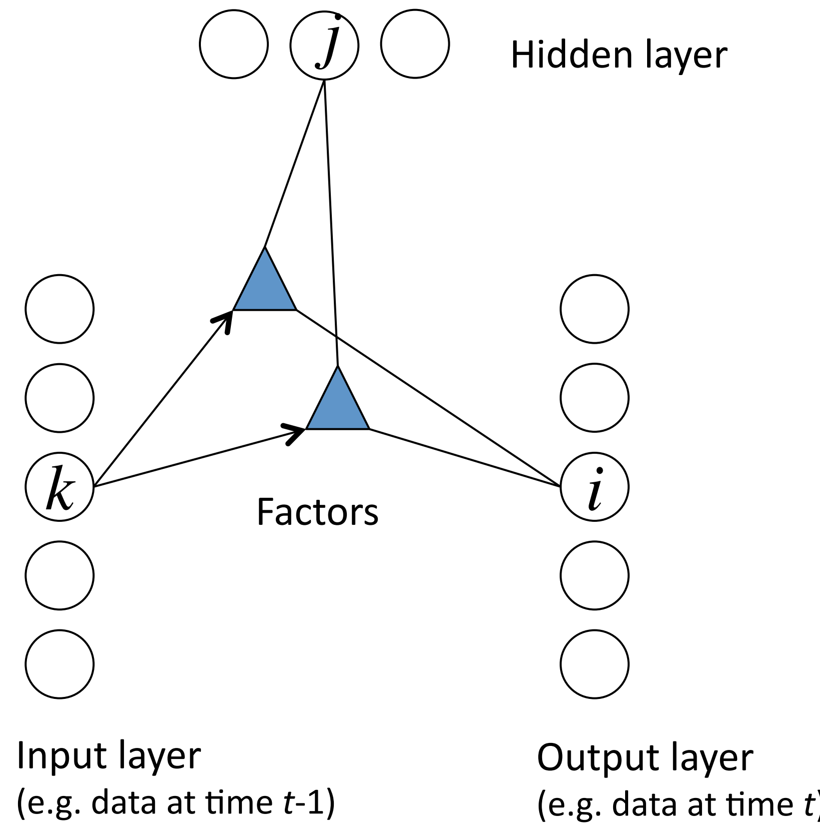




Two equivalent views of Gated Conditional RBMs
(Memisevic and Hinton, 2007)

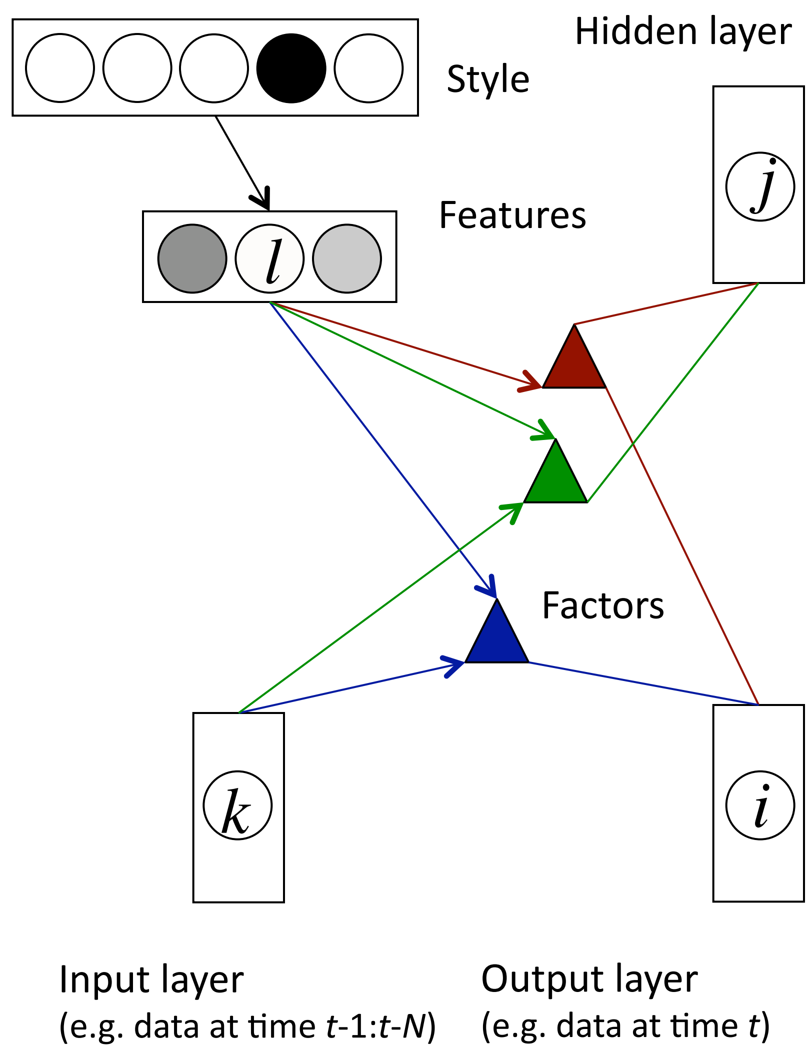
Factoring

- Exponential # of AR models at cubic cost
- Regularities suggest structure can be captured with < cubic # of parameters
- Introduce deterministic "factors": $O(N^3)$ to $O(N^2)$



Factored Conditional RBMs

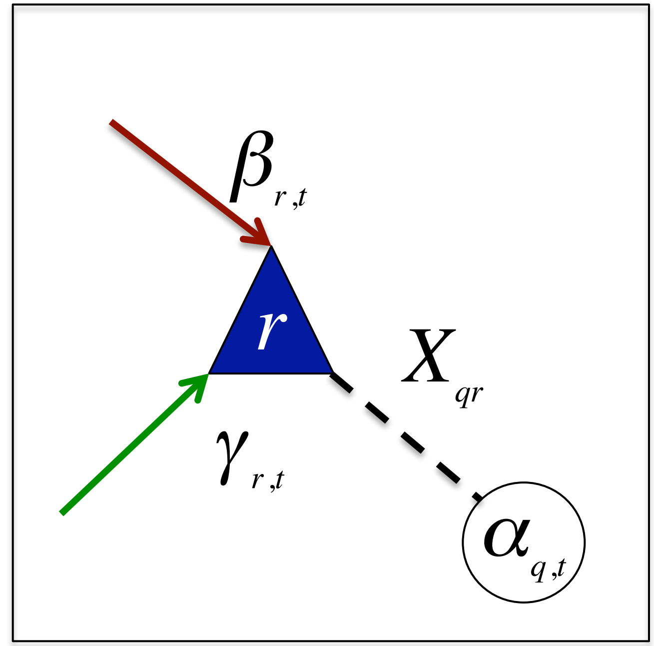
- Let style change *interactions* rather than *effective biases*
- Deterministic features are linearly related to style
- Can blend and transition among motion styles



CD(K) weight updates

Weight updates have the form:

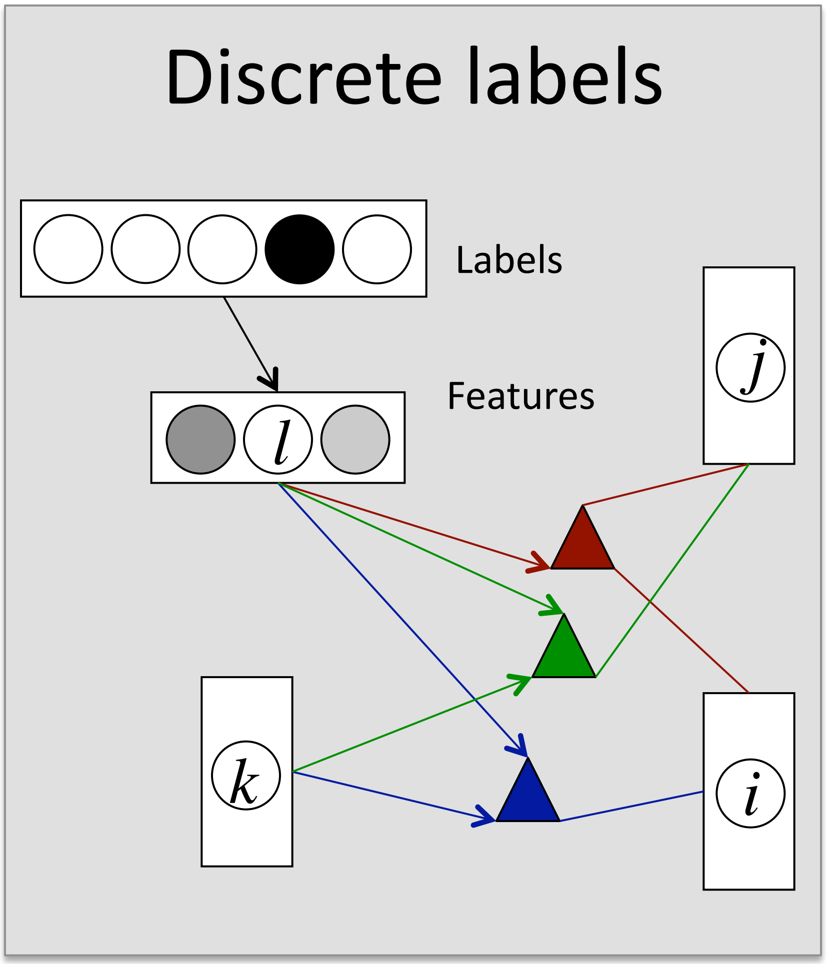
$$\Delta X_{qr} \propto \sum_t \left(\left\langle \alpha_{q,t} \beta_{r,t} \gamma_{r,t} \right\rangle_0 - \left\langle \alpha_{q,t} \beta_{r,t} \gamma_{r,t} \right\rangle_K \right)$$

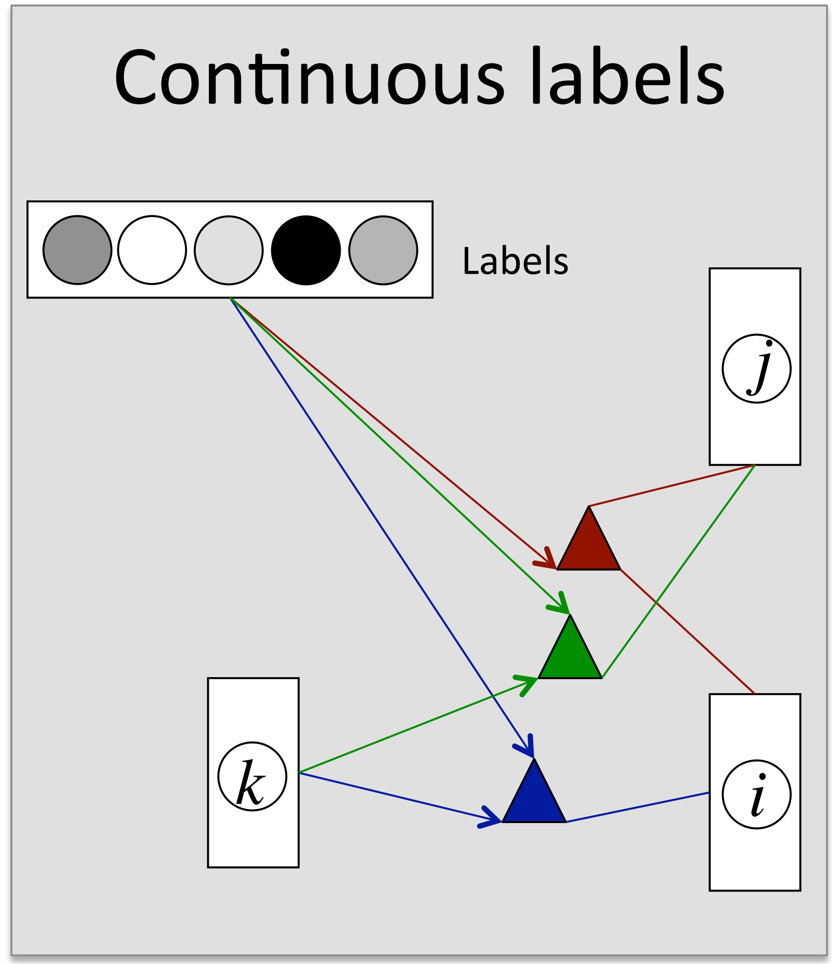


Unit connected to factor r by weight X_{qr}

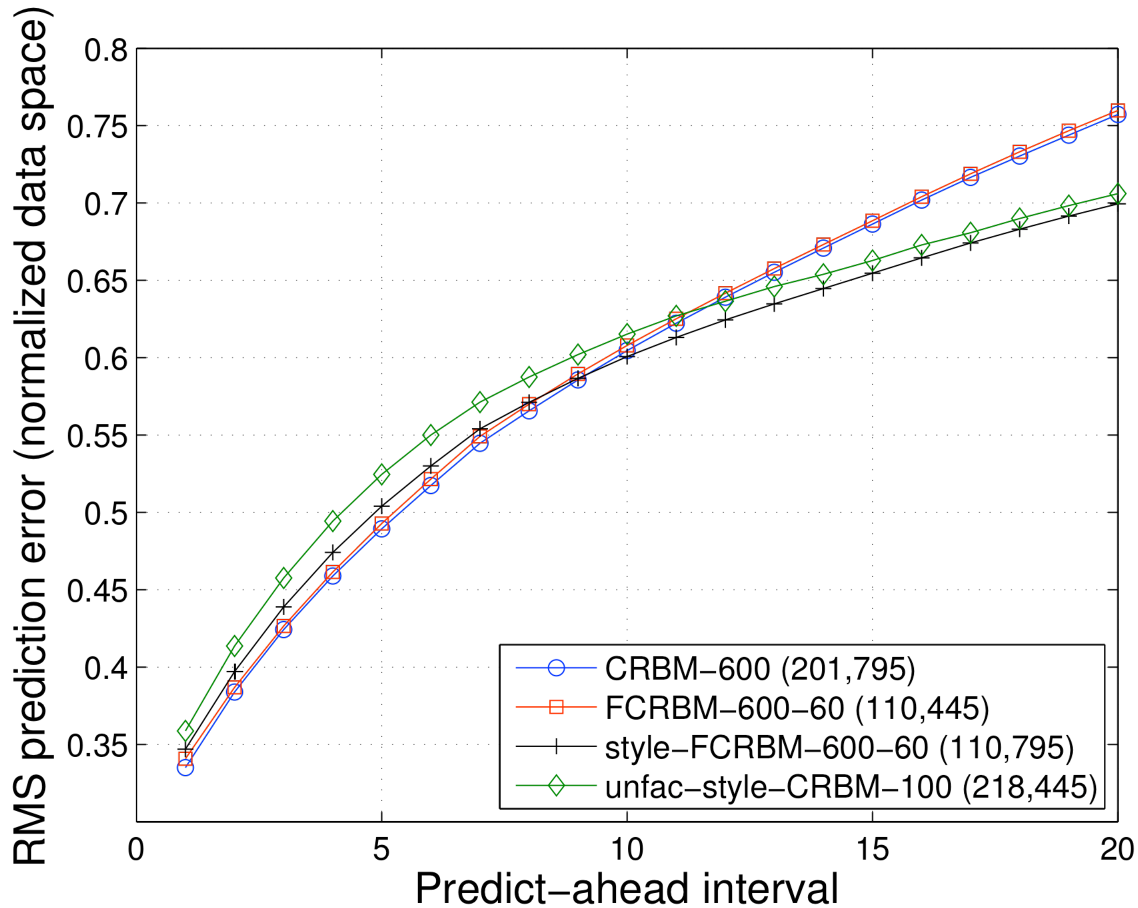
Total input arriving at factor r via each of the two other units involved in the three-way relationship

Results: synthesis





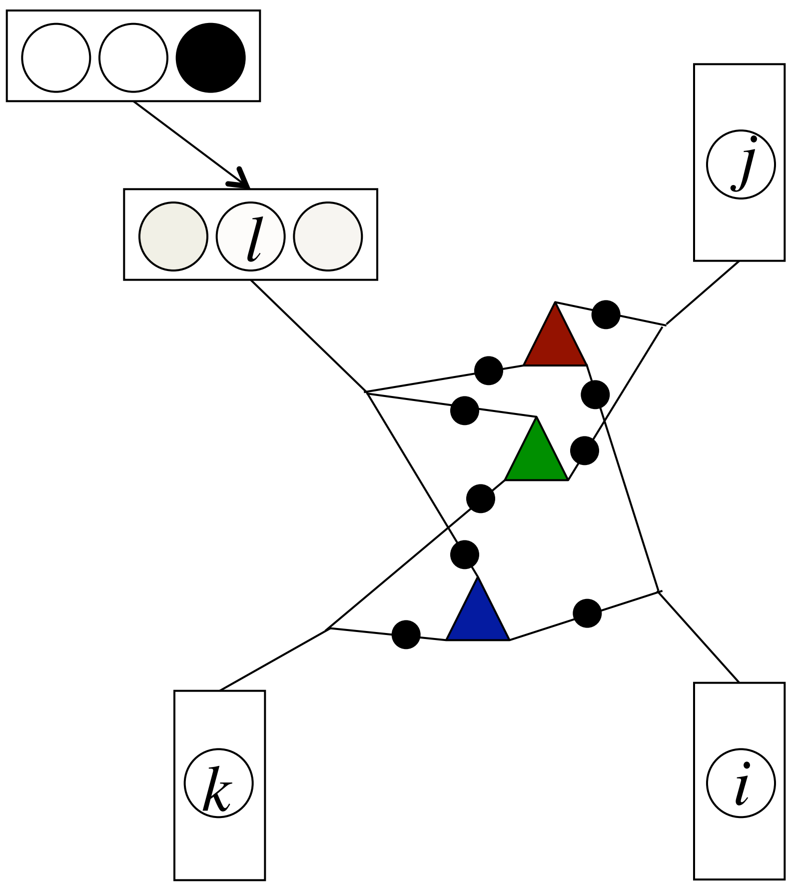
Results: prediction



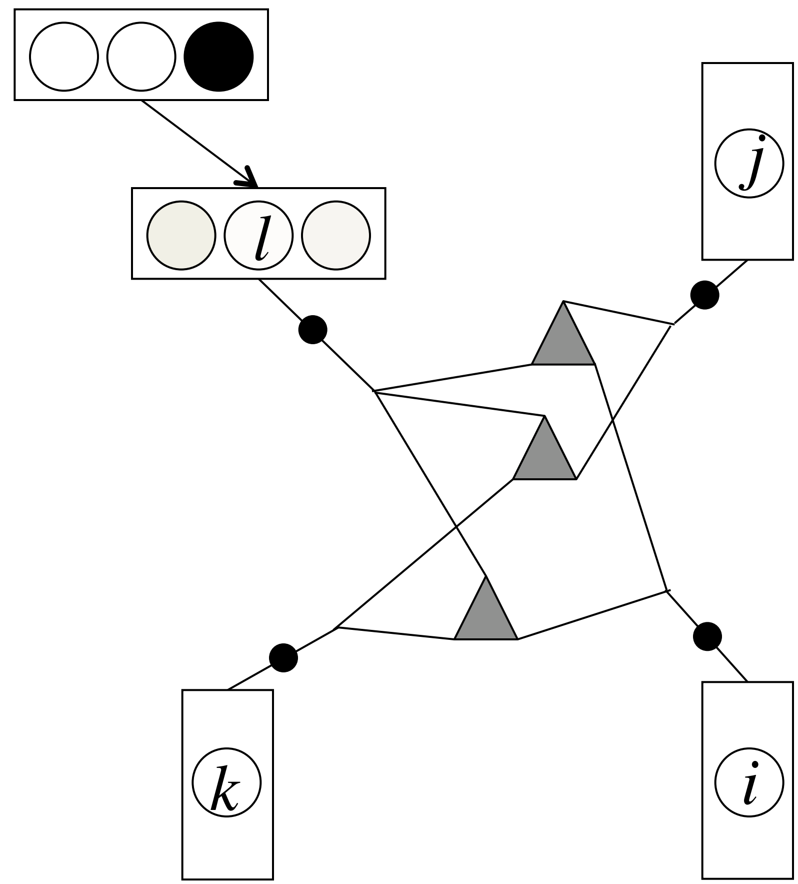
Conclusions

- CRBMs: distributed representations, exact inference and efficient approximate learning
  - Can be composed into conditional DBNs
- FCRBMs: naturally integrate context, multiplicative interactions with quadratic number of parameters
  - Future work: unsupervised style discovery, deep models

Parameter sharing



Fully parameterized
(no sharing)



Full sharing

Related work

Concatenation

Transforming existing motion

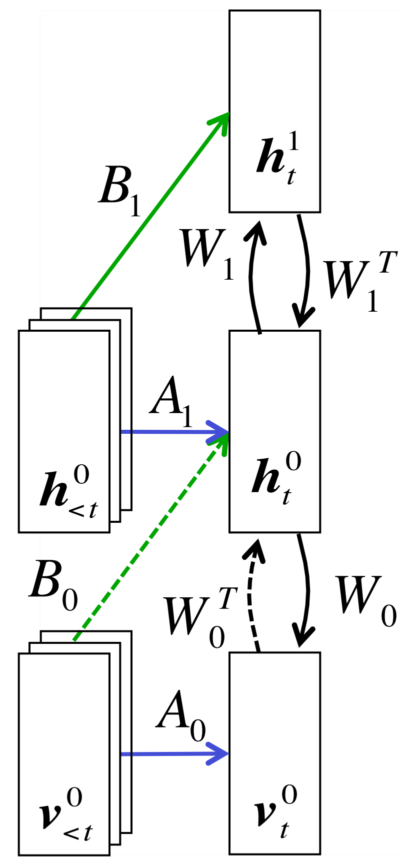
Interpolation/ Blending

Physics-based methods

Generative models-
Our method is based on a "pure" learning approach. It is able to generalize well while avoiding the complexity of explicitly imposing physics-based constraints.

Conditional deep belief networks

- CRBM defines $p(v^0,h^0)$ – implicitly $p(h^0),p(v^0|h^0)$
- Consider "trading in" $p(h^0)$ for a better model
- Subject to conditions (which we violate) – guaranteed to never decrease a variational lower bound on log prob



Modeling human motion

- Capture the movement of a subject as a time series of 3D cartesian coordinates
- High-dimensional (60-100), nonlinear, long-range deps
- Large repositories available

Introduction

- The Conditional Restricted Boltzmann Machine (Taylor et al. 2007)

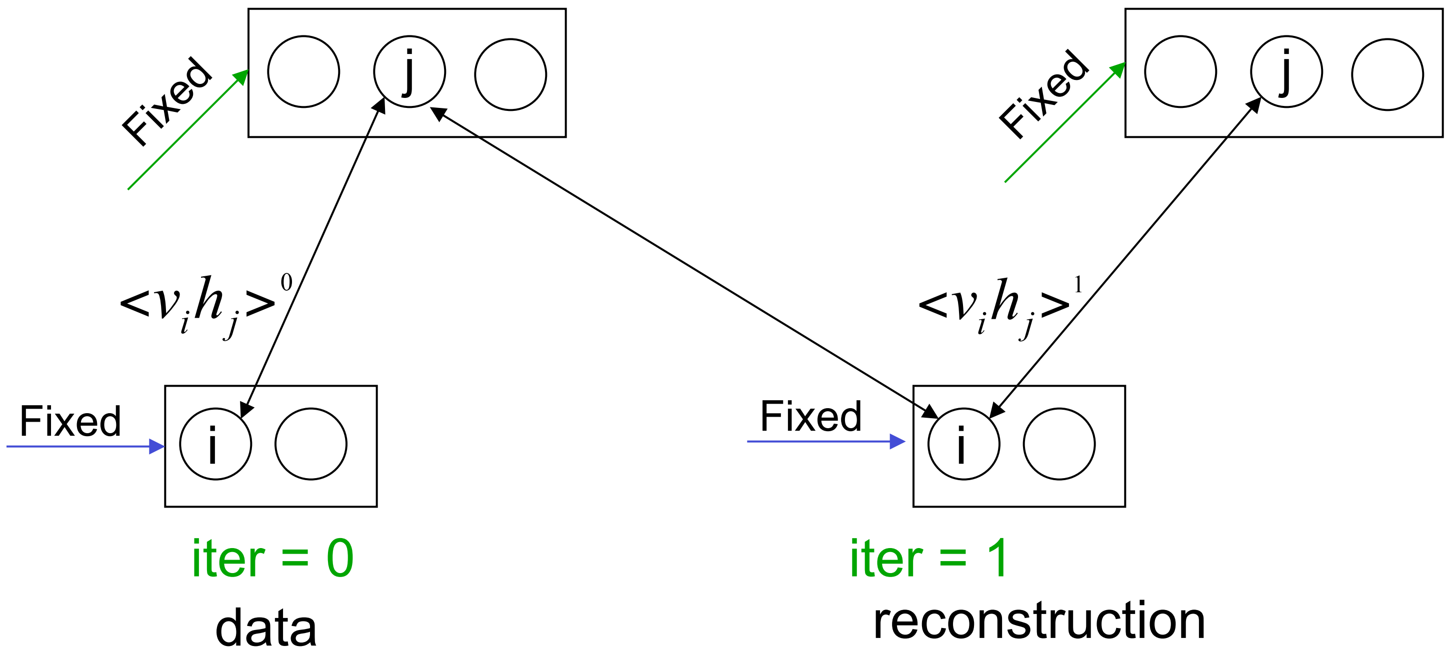
Naturally incorporate contextual information (specifically style) into the CRBM while preserving its most important computational properties.

Joint distribution

$$p(\boldsymbol{v}_t, \boldsymbol{h}_t \mid \boldsymbol{v}_{<t}, \theta)$$

$$= \frac{\exp\left(-E(\boldsymbol{v}_t, \boldsymbol{h}_t \mid \boldsymbol{v}_{<t}, \boldsymbol{y}_t, \theta)\right)}{Z(\theta)}$$

Contrastive divergence learning (CRBM)



When updating visible and hidden units, we implement directed connections by treating data from previous time steps as a dynamically changing bias.

Inference and learning do not change.

Energy function

$$E(\boldsymbol{v}_t, \boldsymbol{h}_t \mid \boldsymbol{v}_{<t}, \boldsymbol{y}_t, \theta) = \frac{1}{2}\sum_i (\hat{a}_{i,t} - v_{i,t})^2$$

$$-\sum_j \hat{b}_{j,t} h_{j,t} - \sum_f \sum_{ijl} W_{if}^{\boldsymbol{v}} W_{jf}^{\boldsymbol{h}} W_{lf}^{\boldsymbol{z}} v_{i,t} h_{j,t} z_{l,t}$$

$$z_{l,t} = \sum_p R_{pl} y_{p,t}$$

$$\hat{a}_{i,t} = a_i + \sum_m A_{im}^{\boldsymbol{v}} \sum_k A_{km}^{\boldsymbol{v}_{<t}} v_{k,<t} \sum_l A_{lm}^{\boldsymbol{z}} z_{l,t}$$

$$\hat{b}_{j,t} = b_j + \sum_n B_{jn}^{\boldsymbol{h}} \sum_k B_{kn}^{\boldsymbol{v}_{<t}} v_{k,<t} \sum_l B_{ln}^{\boldsymbol{z}} z_{l,t}$$