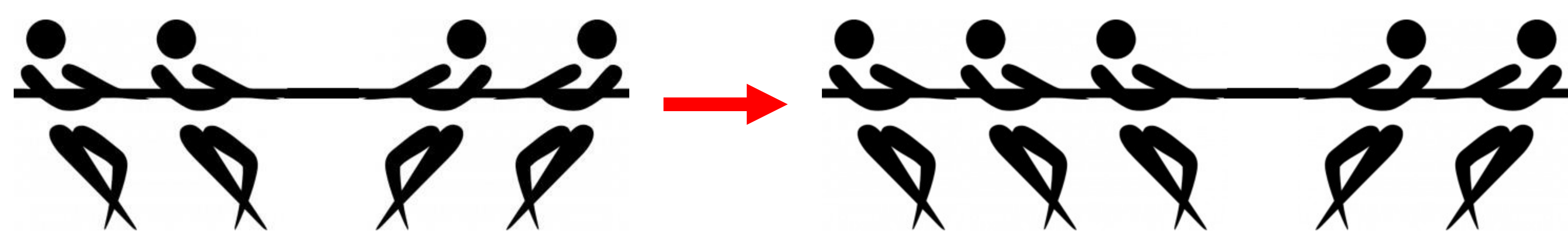


Latent Variable Models (LVM)



Naïve game strategy: Maximize your score

LVM Scoring Function

$$S_{LVM}(x, y) = \max_{a \in Z} f_{\theta_y}(x, a)$$

Labels: input (x), class (y), latent variable (a), model parameters (θ_y)

LVM looks for *evidences* of a class

A Simple Idea!

"cow" Image classification: different latent variables



Generalized Latent Variable Models (GLVM)



Better game strategy: Maximize your score and minimize your opponent's

GLVM Scoring Function

$$S_{GLVM}(x, y) = \max_{a \in Z^+} f_{\theta_y}(x, a) - \max_{b \in Z^-} f_{\theta_y}(x, b)$$

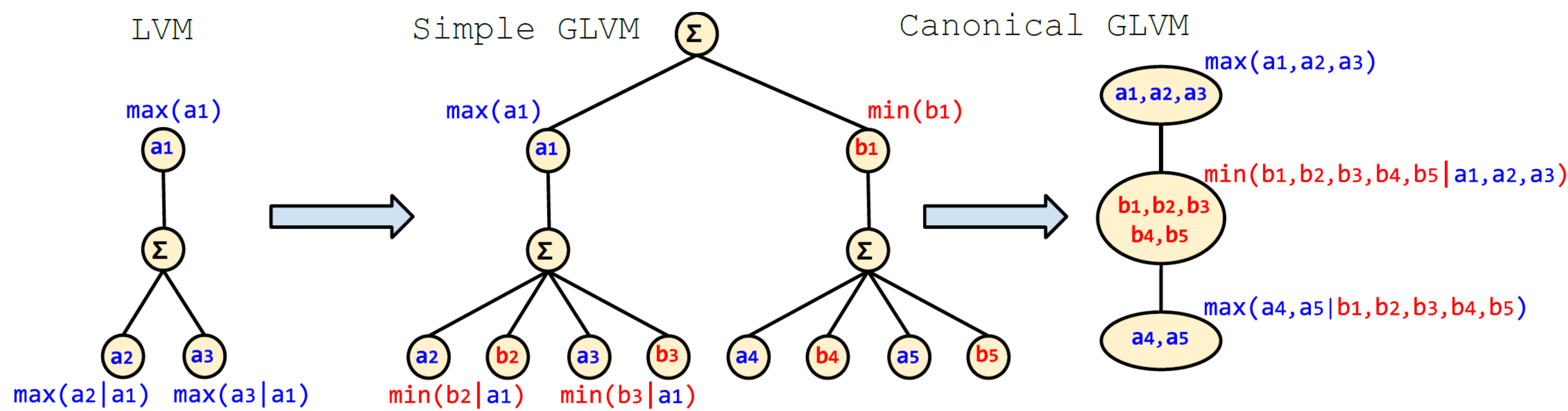
Labels: positive latent variable (a), negative latent variable (b)

GLVM looks for both *evidences* and *counter evidences* of a class

Recursive positive and negative latent variables



Dependency Structure of Latent Variables



Connections

- Deformable Part Models can be generalized
- Latent Hough Transform can be generalized
- And-Or trees can be generalized
- and probably many more discriminative LVMs
- Latent Structural SVMs as a shallow case of GLVM [2]
- Mid-level features for scene recognition [1,3,4]
- ConvNets have negative parts!? [5]

Cat Head Detection (Oxford Pet)

	Abyssinian	Bengal	Birman	Bombay	British Shorthair	Egyptian Mau	Maine Coon	Persian	Ragdoll	Russian Blue	Siamese	Sphynx
DPM_4	21.3	12.8	34.5	23.3	32.2	15.8	21.6	28.0	19.4	24.0	29.4	22.7
DPM_6	22.2	12.8	31.4	21.5	31.3	16.5	26.0	29.0	20.6	25.0	30.9	22.0
$GDPM_2^2$	24.3	11.9	38.4	23.9	31.0	19.7	27.7	29.7	24.5	29.9	35.9	27.4
$GDPM_4^2$	25.5	13.7	34.0	23.0	33.5	20.9	24.5	30.0	21.9	25.3	30.6	23.2

- Generalized DPM implementation is non-trivial
- Consistent improvement
- No meaningful visualization of negative parts
- More robustness toward over-fitting

PASCAL VOC 2007 Animals

	bird	cat	cow	dog	horse	sheep
DPM_8	10.9	16.6	21.9	12.5	56.0	18.0
$GDPM_7^1$	10.9	17.0	22.0	12.5	57.2	19.2
$GDPM_6^2$	9.5	18.6	23.2	12.2	56.0	19.8
$GDPM_5^3$	10.5	13.9	21.7	11.9	54.8	19.1

Initialization is important:

AP for careful initialization of $GDPM_8^1$ for cow is **26.9**

