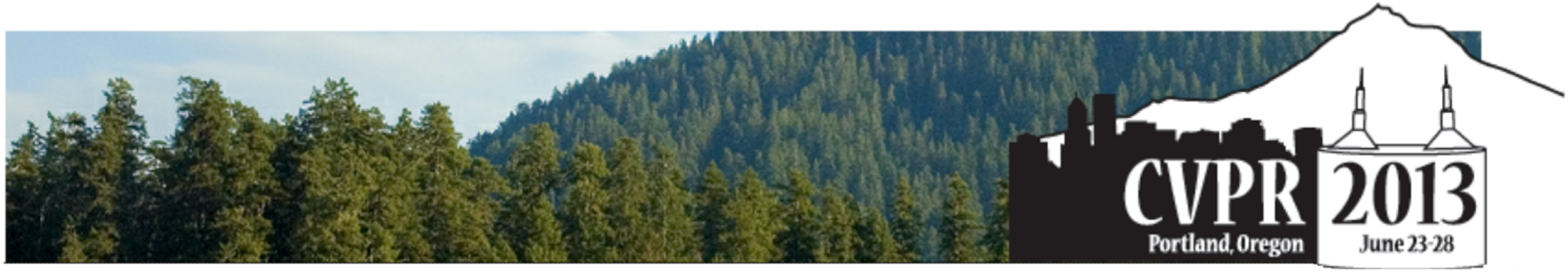


Attributes

Devi Parikh
Virginia Tech

Ali Farhadi
UW

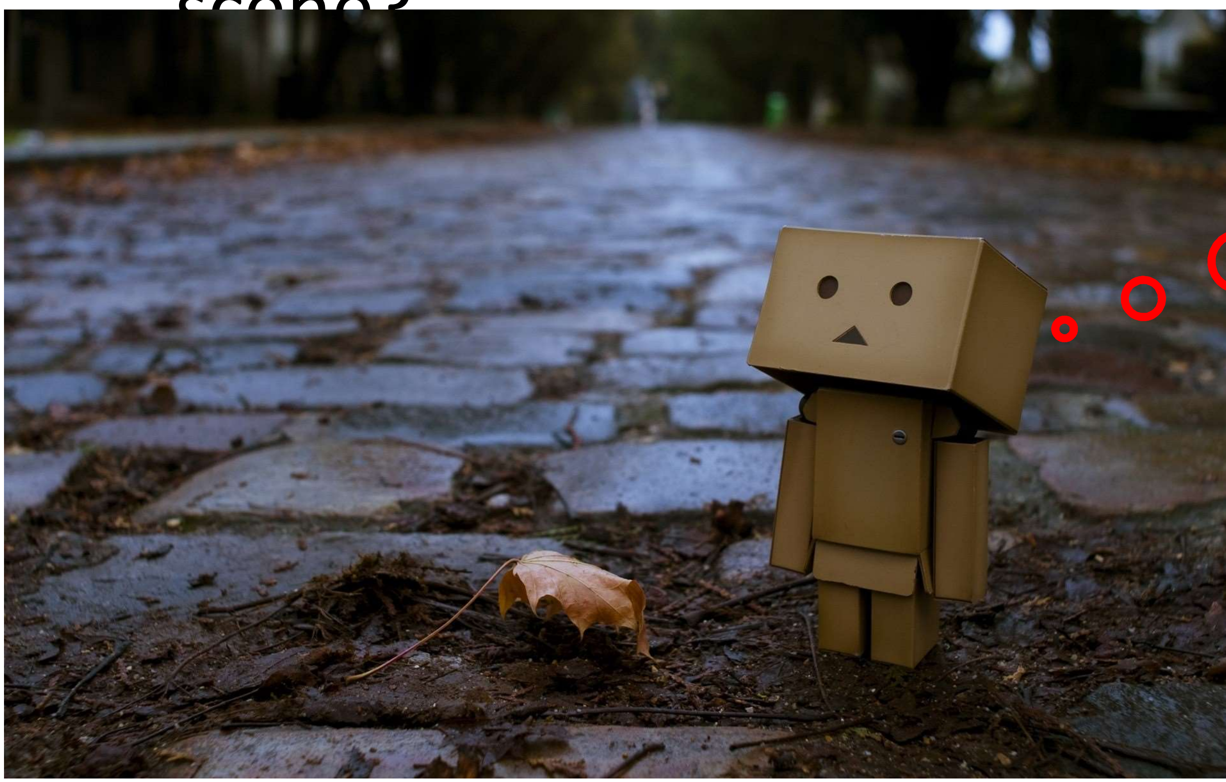
Kristen Grauman
UT Austin

Tamara Berg
UNC (Fall 13)

Abhinav Gupta
CMU

Why attributes?

- Why recognition?
- Why would a robot need to recognize a scene?



Can I walk around here? Is this walkable?

Slide credit: Devi Parikh

Why attributes?

- Why would a robot need to recognize an object?

How hard should I grip this? Is it brittle?

Slide credit: Devi Parikh

Why attributes?

- How humans naturally describe visual concepts
- Image search

I want elegant silver sandals with high heels

Slide credit: Devi Parikh

Why attributes?

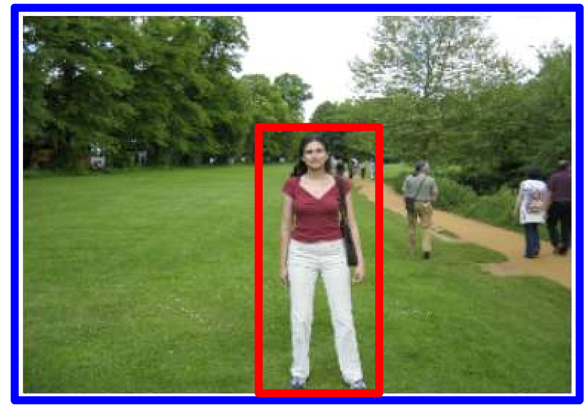
- How humans think of concepts
- Domain knowledge helpful to build visual models

Slide credit: Devi Parikh

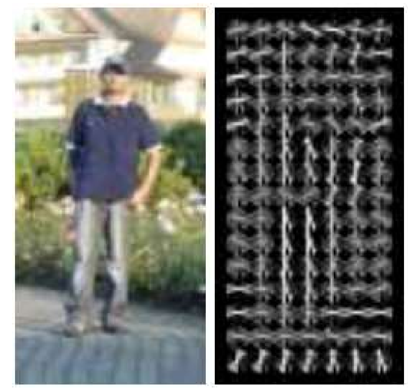
What are attributes?

- Mid-level concepts
  - Higher than low-level features
  - Lower than high-level categories
- Shared across categories
- Human-understandable (semantic)
- Machine-detectable (visual)

Outdoor, Peaceful, Natural, Green,…



Female, Long-hair, Young …



Dalal and Triggs, 2005

Ladicky et al, 2010

Slide credit: Devi Parikh

What are attributes?

- Material, Appearance, Function, …
- Any adjective
- Statements *about* visual concepts


¹at·tri·bute *noun* \'a-trə-ˌbyüt\

Definition of ATTRIBUTE

1 : an inherent characteristic; *also* : an accidental quality

2 : an object closely associated with or belonging to a specific person, thing, or office <a scepter is the *attribute* of power>; *especially* : such an object used for identification in painting or sculpture

3 : a word ascribing a quality; *especially* : ADJECTIVE

- Objects, scene: Nouns
- Actions: Verbs

- Distinctions less critical in how to predict attributes
- Distinctions important in uses of attributes for improving computer vision

- Parts often semantic, shared, mid-level
- Can be used as attributes

Slide credit: Devi Parikh

Example Attributes

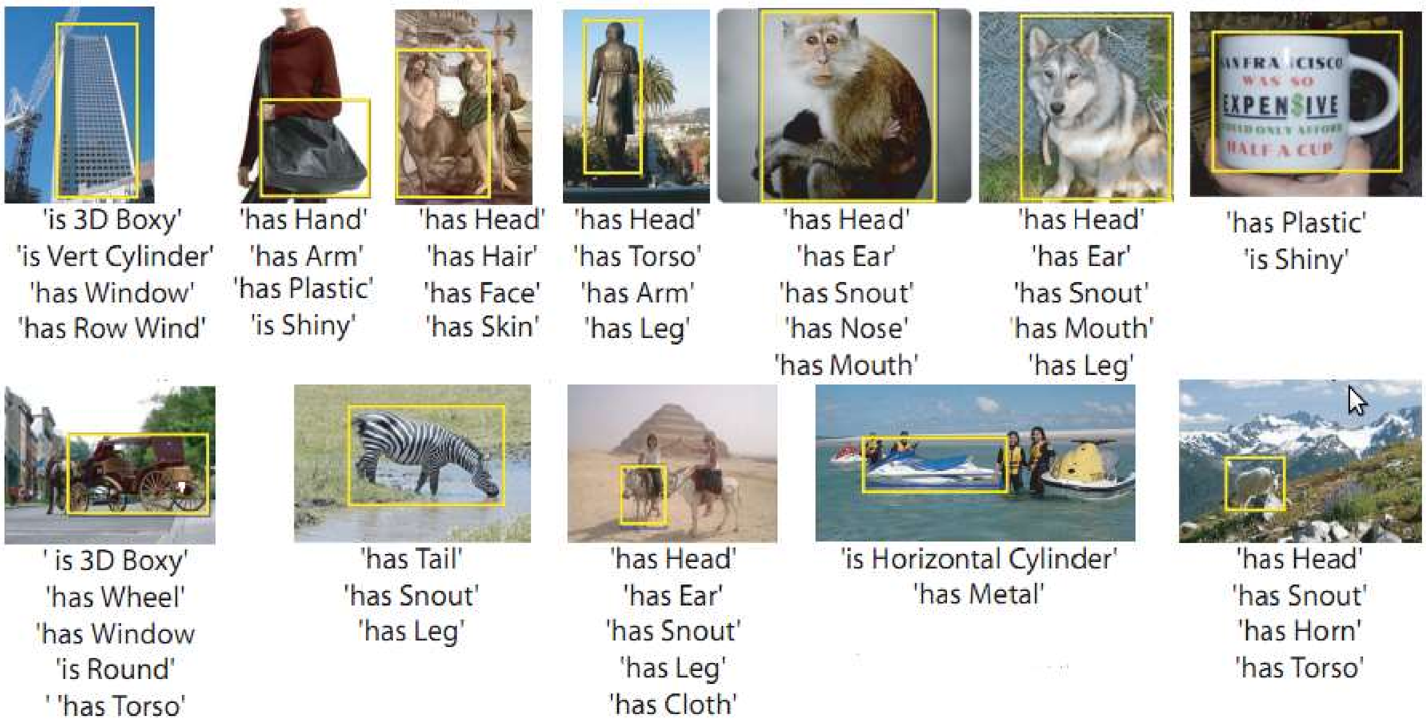
Face Tracer Image Search

“Smiling Asian Men With Glasses”

Kumar et al. 2008

Slide credit: Devi Parikh

Example Attributes



Farhadi et al. 2009

Slide credit: Devi Parikh

Example Attributes

<u>otter</u>

black: yes
white: no
brown: yes
stripes: no
water: yes
eats fish: yes

<u>polar bear</u>

black: no
white: yes
brown: no
stripes: no
water: yes
eats fish: yes

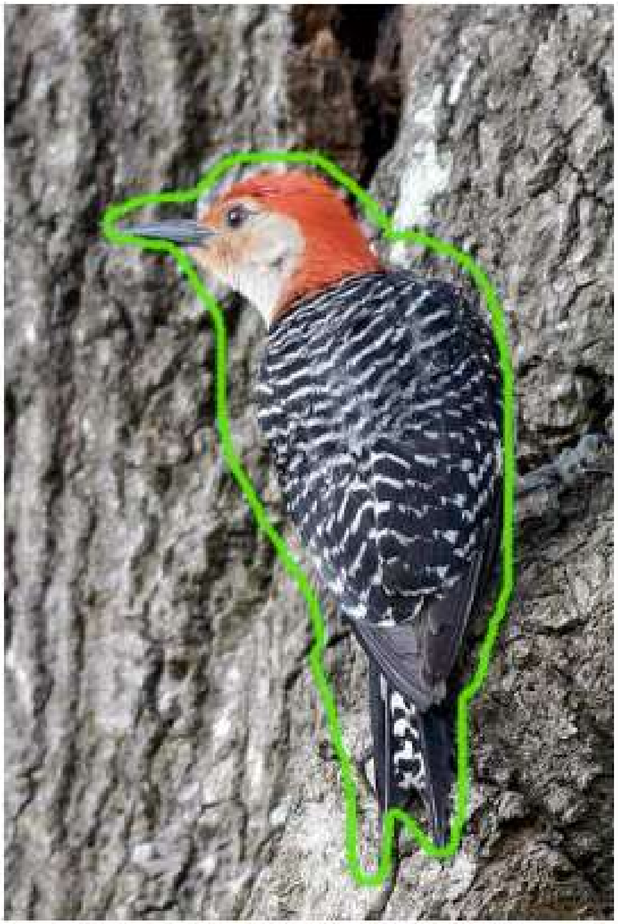
<u>zebra</u>

black: yes
white: yes
brown: no
stripes: yes
water: no
eats fish: no

Lampert et al. 2009

Slide credit: Devi Parikh

Example Attributes

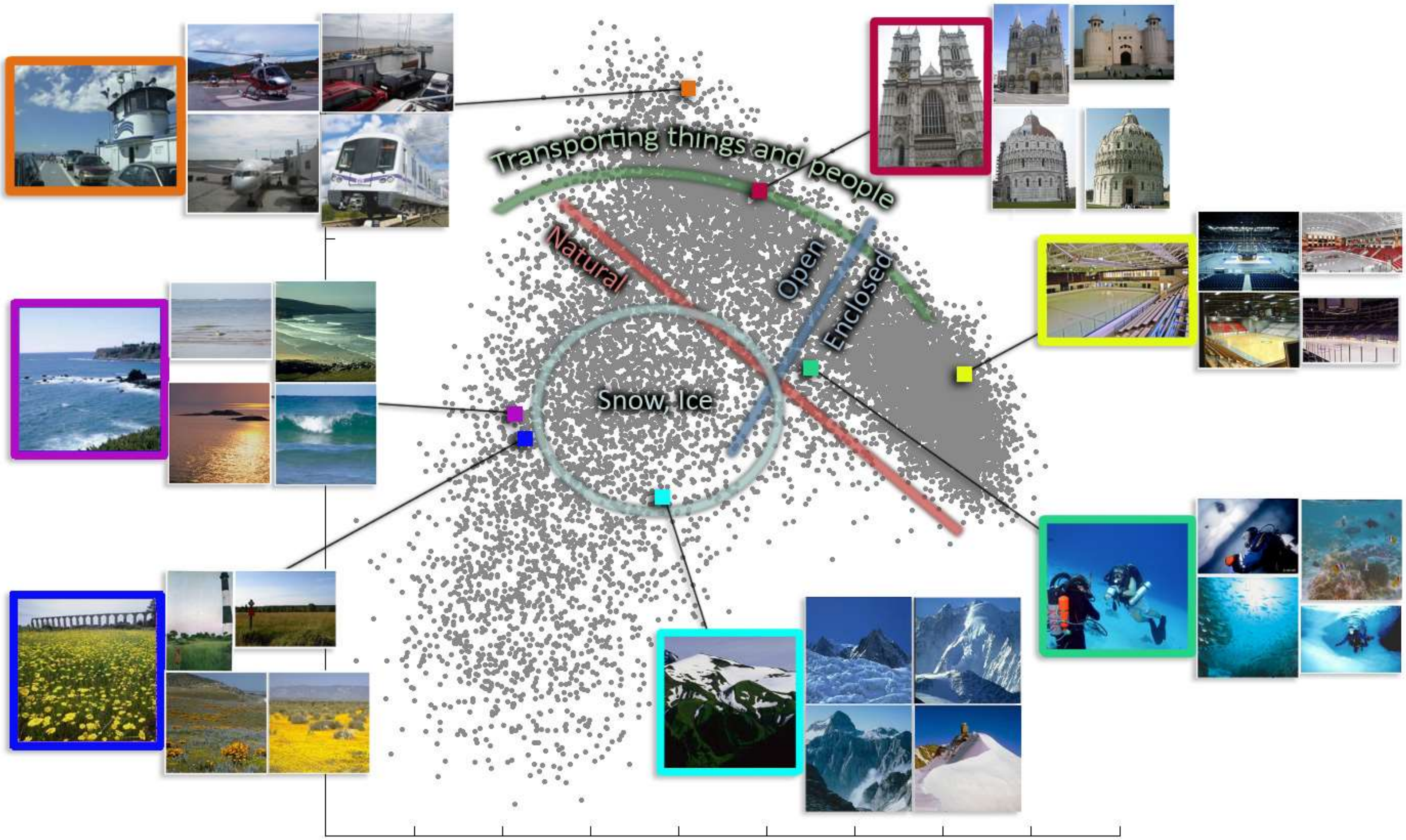


forehead_color	**red**	red	**red**
breast_pattern	multi-colored	solid	solid
breast_color	white	white/red	white
head_pattern	**capped**	capped	**capped**
back_color	**white/ black**	**white/ black**	**white/ black**
wing_color	**white/ black**	white/black	white/black
leg_color	buff	black	black
size	**small**	medium	medium
bill_shape	**all-purpose**	**dagger**	**all-purpose**
wing_shape	**pointed**	tapered	pointed
...	...	...	...
primary_color	black, red	white, black	**white, black**

Welinder et al. 2010

Slide credit: Devi Parikh

Example Attributes



Patterson and Hays 2011

Slide credit: Devi Parikh

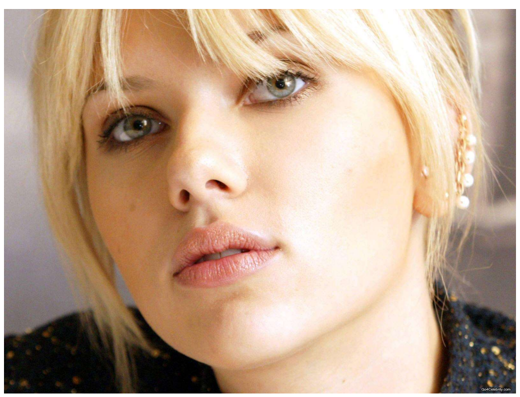
Example Attributes

Berg et al. 2010

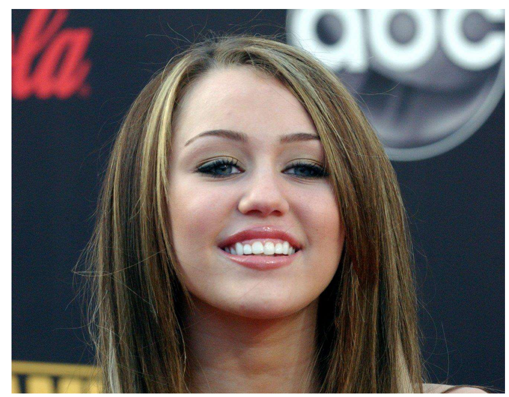
Slide credit: Devi Parikh

Relative Attributes

>
natural



<
smiling



Parikh and Grauman 2011

Slide credit: Devi Parikh

Some Notation

- Vocabulary of attributes:

$$A = \{a_m\}, m \in \{1, \ldots, M\}$$

- Image features:

$$\{\boldsymbol{x}_i\}, i \in \{1, \ldots, N\}$$

Slide credit: Devi Parikh

Attribute Models

- Classifiers for binary attributes

$$x_i \rightarrow \{+1, -1\}$$

(Or confidence)

Kumar et al. 2010

Slide credit: Devi Parikh

Attribute Models

Noisy labels !

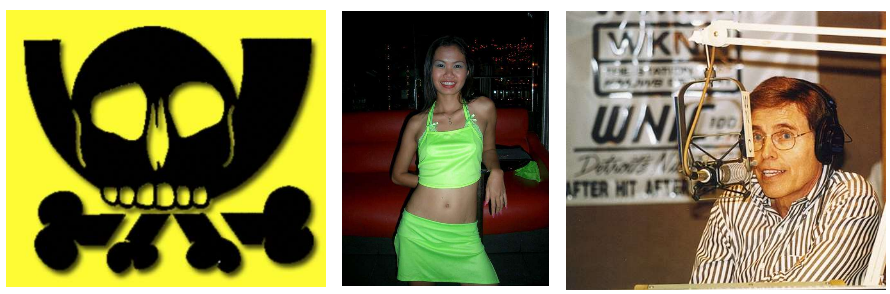
Weakly supervised learning

"stripes"

Positive training images
(unsegmented)

Attribute model

"things"



Negative training images

Ferrari et al. 2007

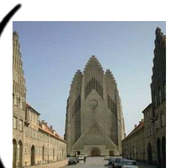
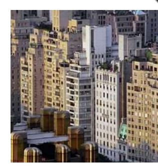
Slide adapted from Vittorio Ferrari by Devi Parikh

Attribute Models

- Ranking functions for relative attributes

For each attribute a_m, open

Supervision is

$O_m : \{($  $\succ$  $), \dots\}$,

$S_m : \{\{$  $\sim$  $\}, \dots\}$

Slide credit: Devi Parikh

Attribute Models

Learn a scoring function $r_m(\boldsymbol{x_i}) = \boldsymbol{w_m^T x_i}$

Image features

Learned parameters

that best satisfies constraints:

$$\forall (i,j) \in O_m : \boldsymbol{w_m^T x_i} > \boldsymbol{w_m^T x_j}$$

$$\forall (i,j) \in S_m : \boldsymbol{w_m^T x_i} = \boldsymbol{w_m^T x_j}$$

Max-margin learning to rank formulation of Joachims 2002

Slide credit: Devi Parikh

Attribute Models

$x_i \rightarrow$ Real value
(strength of attribute presence)

Density

"I am 60% sure this person is smiling"
(Binary Classifier Confidence)

"This person is smiling 60%"
(Attribute Strength)

"Person A is smiling more than Person B"
(Relative Attribute)

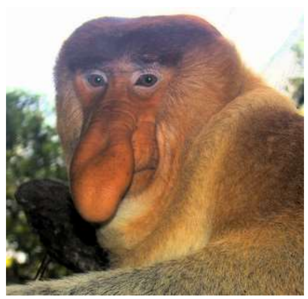
Slide credit: Devi Parikh

Offline Uses of Attributes

Zero-shot Learning

- Aye-ayes
  - Are nocturnal
  - Live in trees
  - Have large eyes
  - Have long middle fingers

Which one of these is an aye-aye?

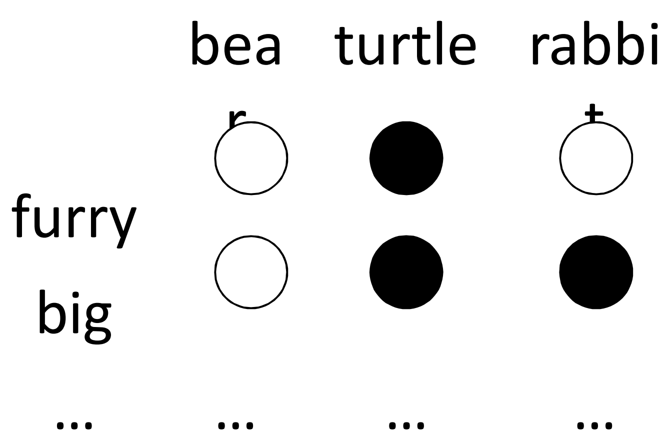



Humans can learn from descriptions (zero examples).

Slide adapted from Christoph Lampert by Devi Parikh

Zero-shot Learning

- Seen categories with labeled images
  - Train attribute predictors
- Unseen categories
  - No examples, only description



Slide credit: Devi Parikh

Zero-shot Learning

- Test image: $\boldsymbol{x}$
- Test class: z
- Classification: $\operatorname{argmax}_z \quad p(z|\boldsymbol{x})$

Zero-shot Learning

Lampert et al. 2009

$p(z|\boldsymbol{x})$

$z =$ zebra

$\boldsymbol{a} = \begin{bmatrix} \text{striped } 1 \\ \text{4-legs } 1 \\ \text{large} \\ 0 \end{bmatrix}$
...

z is independent of $\boldsymbol{x}$ given $\boldsymbol{a}$

0 for all other a

prior (e.g. uniform)

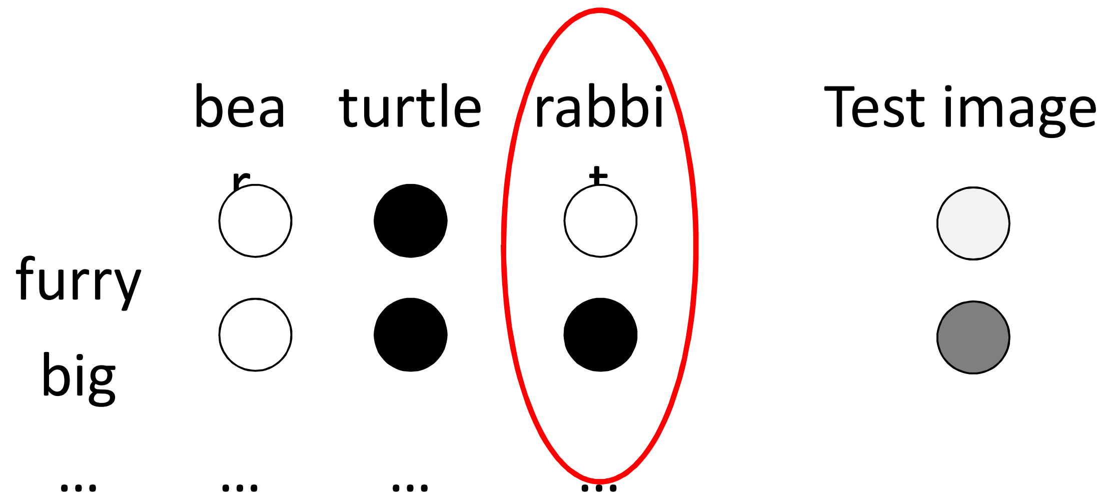
estimate from data or uniform

(assuming uniform priors above)

outputs of binary classifiers

Slide credit: Devi Parikh

Zero-shot Learning



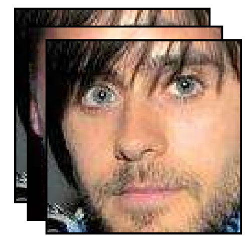
Farhadi et al. 2009

Relative Zero-shot Learning

Training: Images from **S seen** categories and

Descriptions of **U unseen** categories

Age: **Hugh** $\succ$ **Clive** $\succ$ **Scarlett** **Jared** $\succ$ **Miley**



Smiling: **Miley** $\succ$ **Jared**

Need not use all attributes, or all seen categories

Testing: Categorize image into one of **S+U** categories

Slide credit: Devi Parikh

Relative Zero-shot Learning

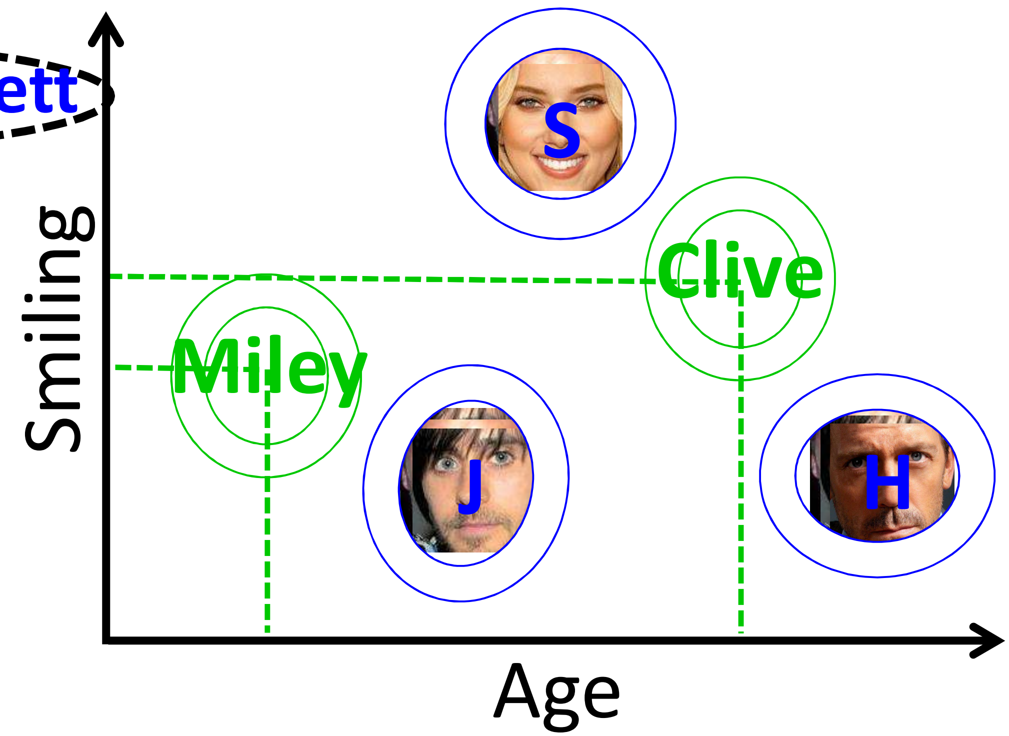
Parikh and Grauman 2011

Can predict new classes based on their relationships to existing classes – without training images

Age: Hugh $\succ$ Clive $\succ$ Scarlett

Jared $\succ$ Miley

Smiling: Miley $\succ$ Jared

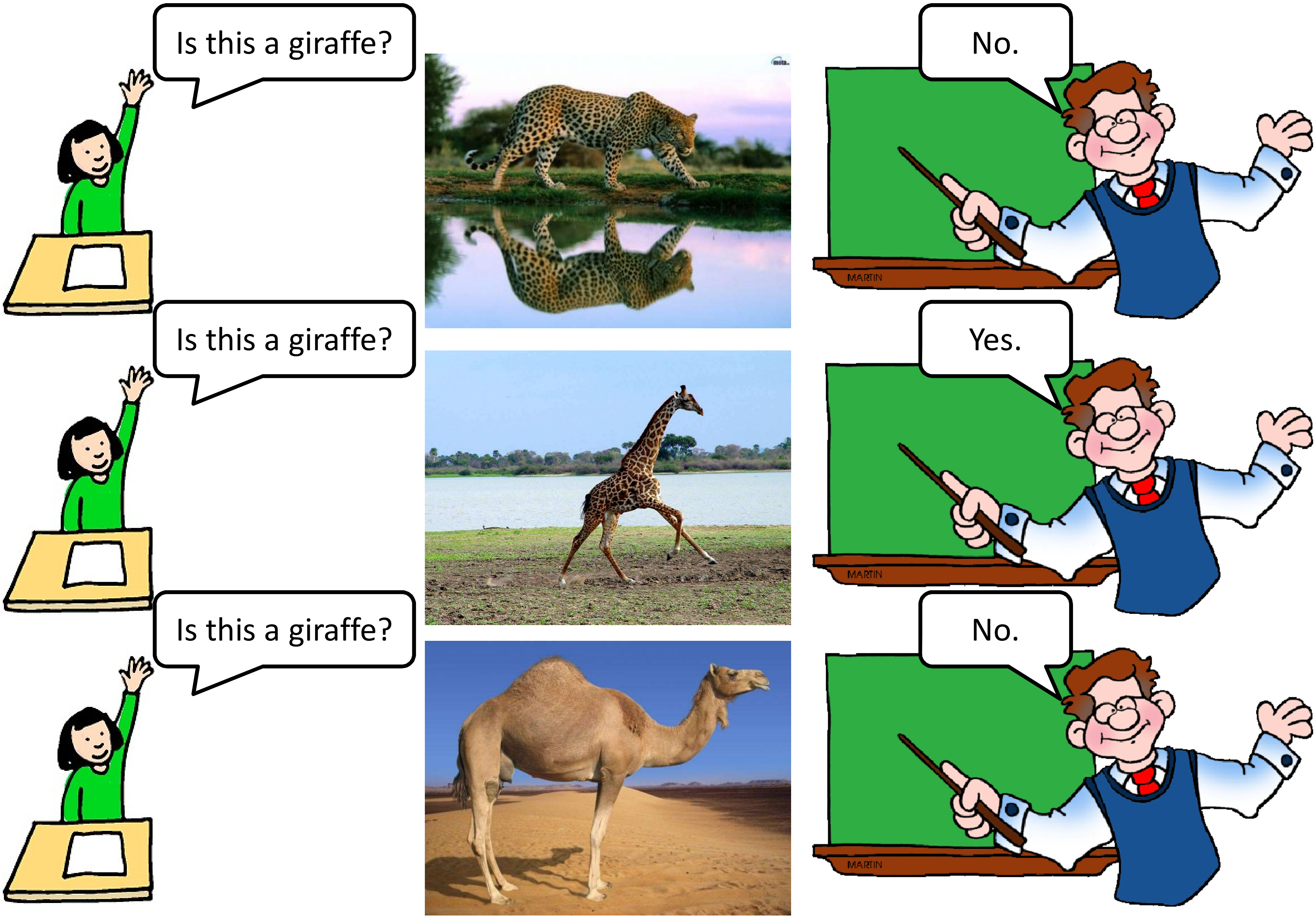


Infer image category using max-likelihood

Slide credit: Devi Parikh

Learning with a few examples

- Active Learning



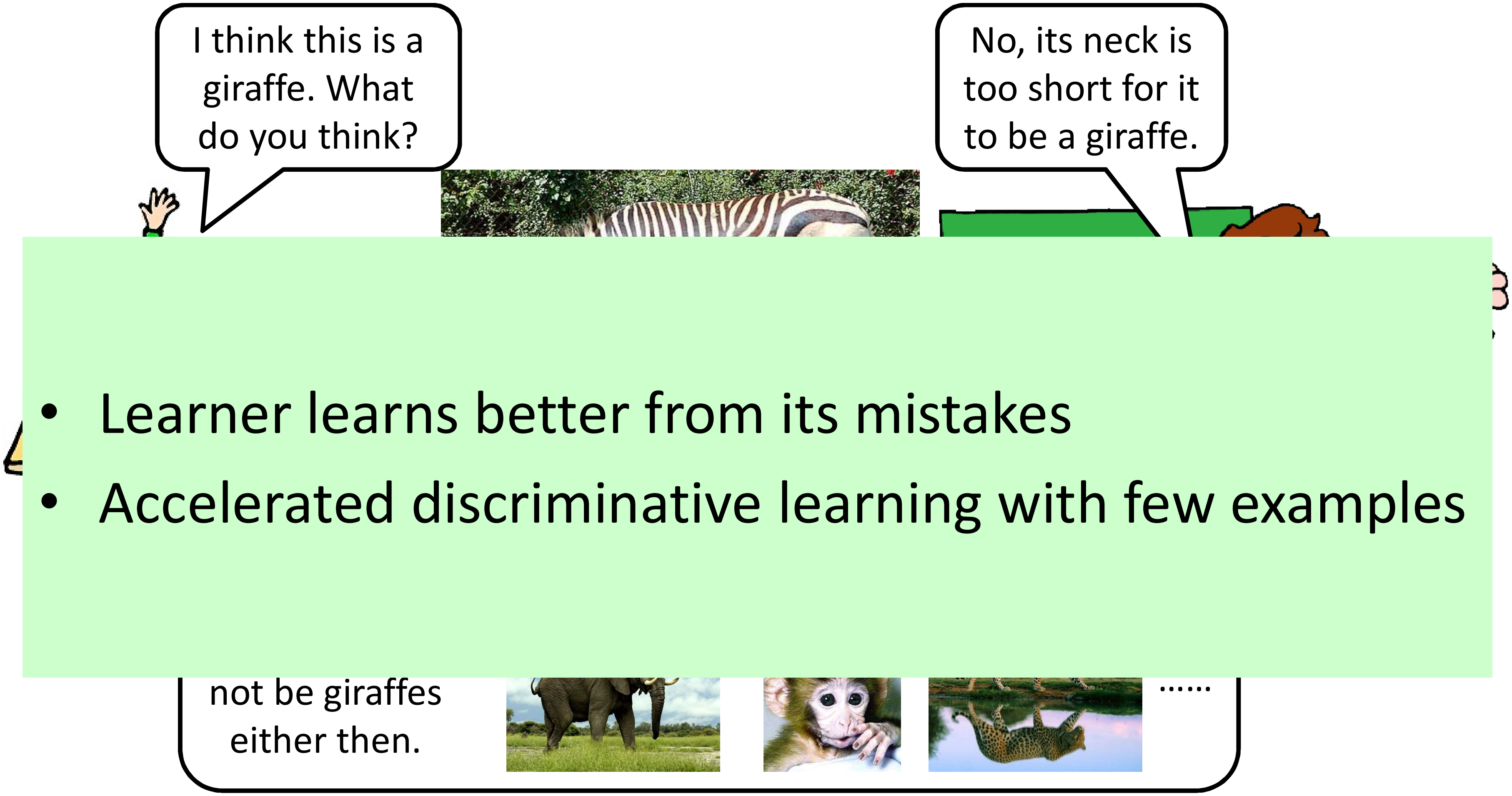
Slide credit: Devi Parikh

Parkash and Parikh, 2012

Current belief

Focused feedback

Knowledge of the world

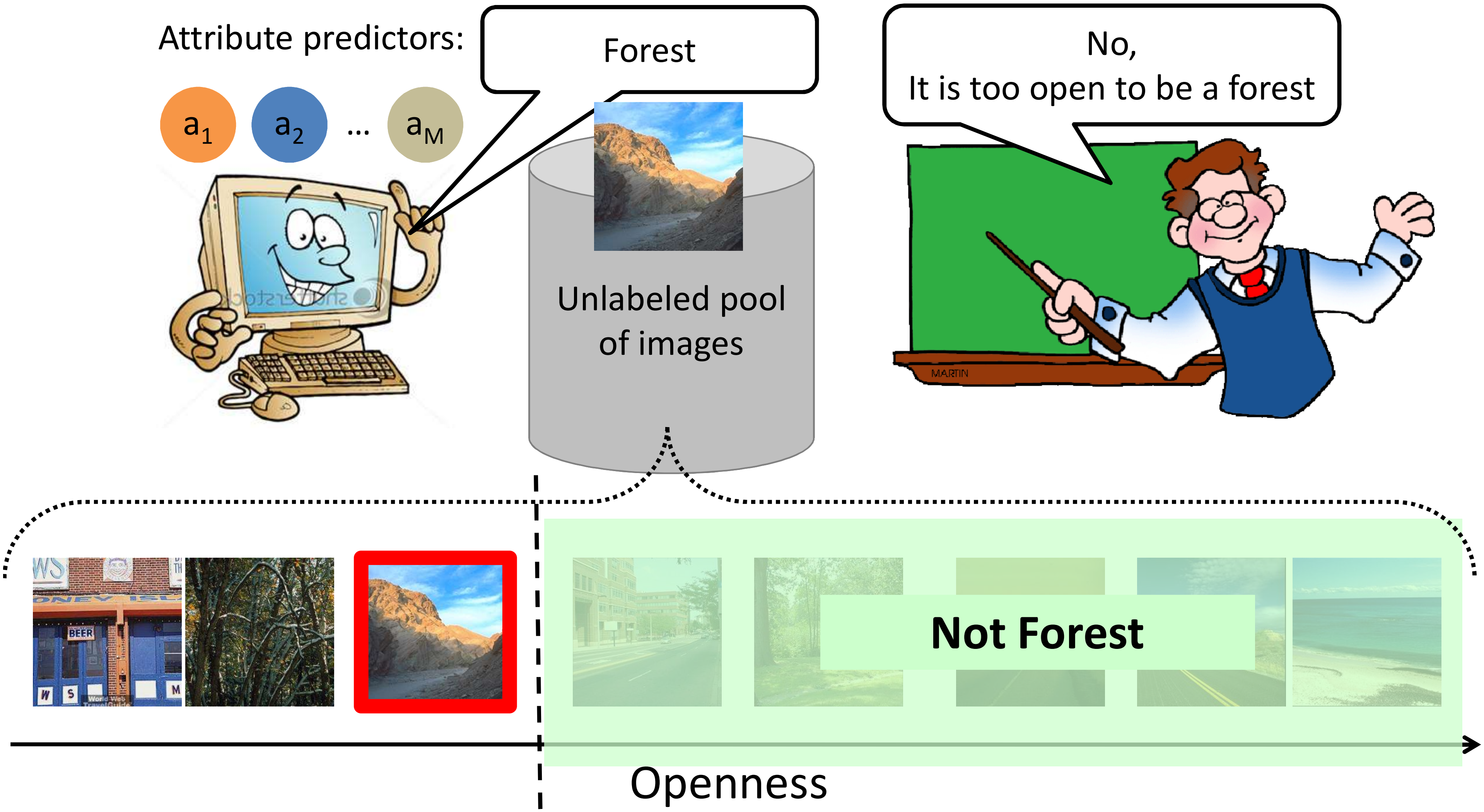


- Learner learns better from its mistakes
- Accelerated discriminative learning with few examples

Feedback on one, transferred to many

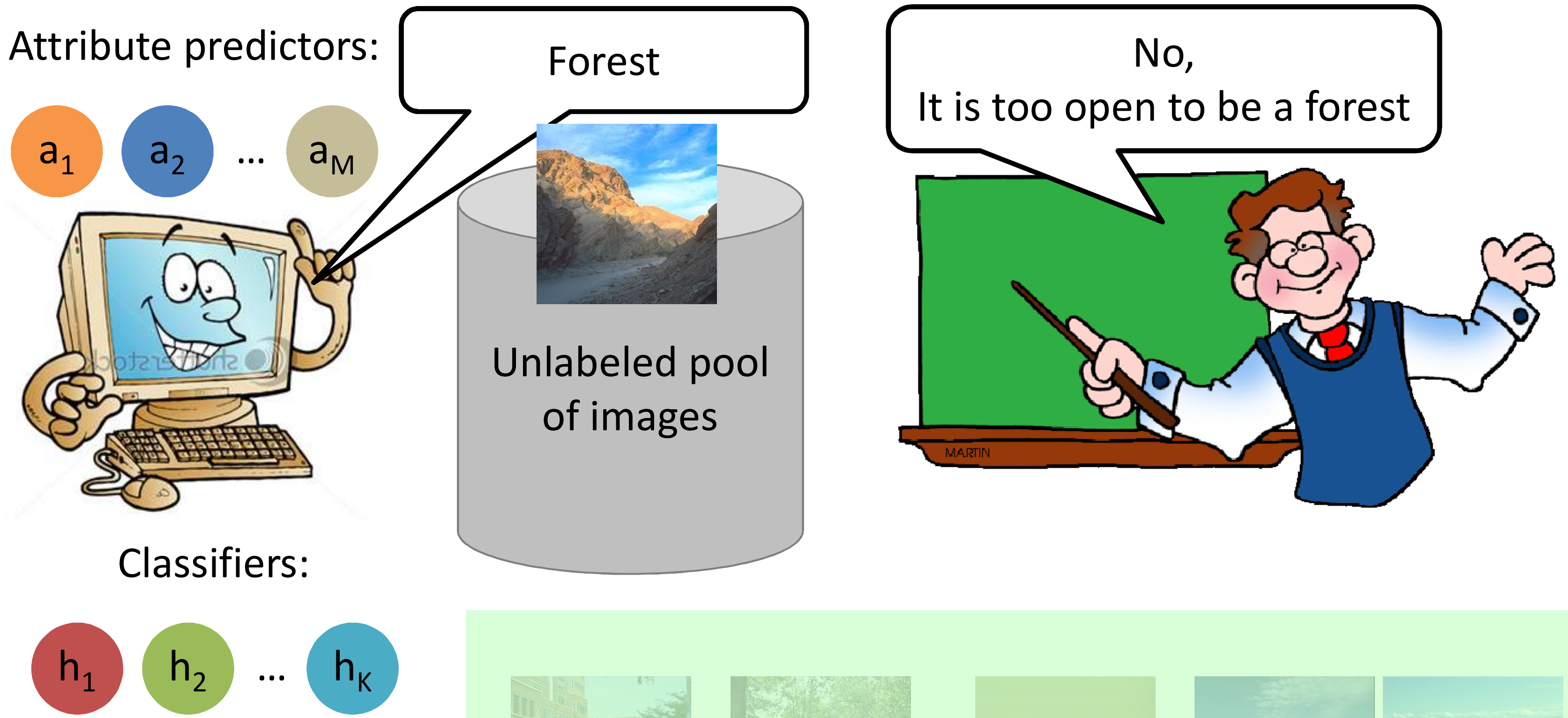
Slide credit: Devi Parikh

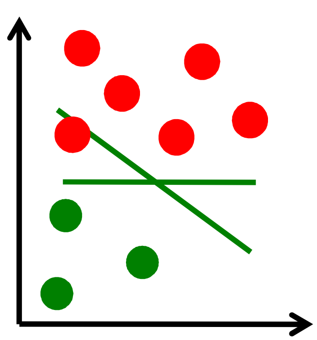
Attributes-based Feedback



Slide credit: Devi Parikh

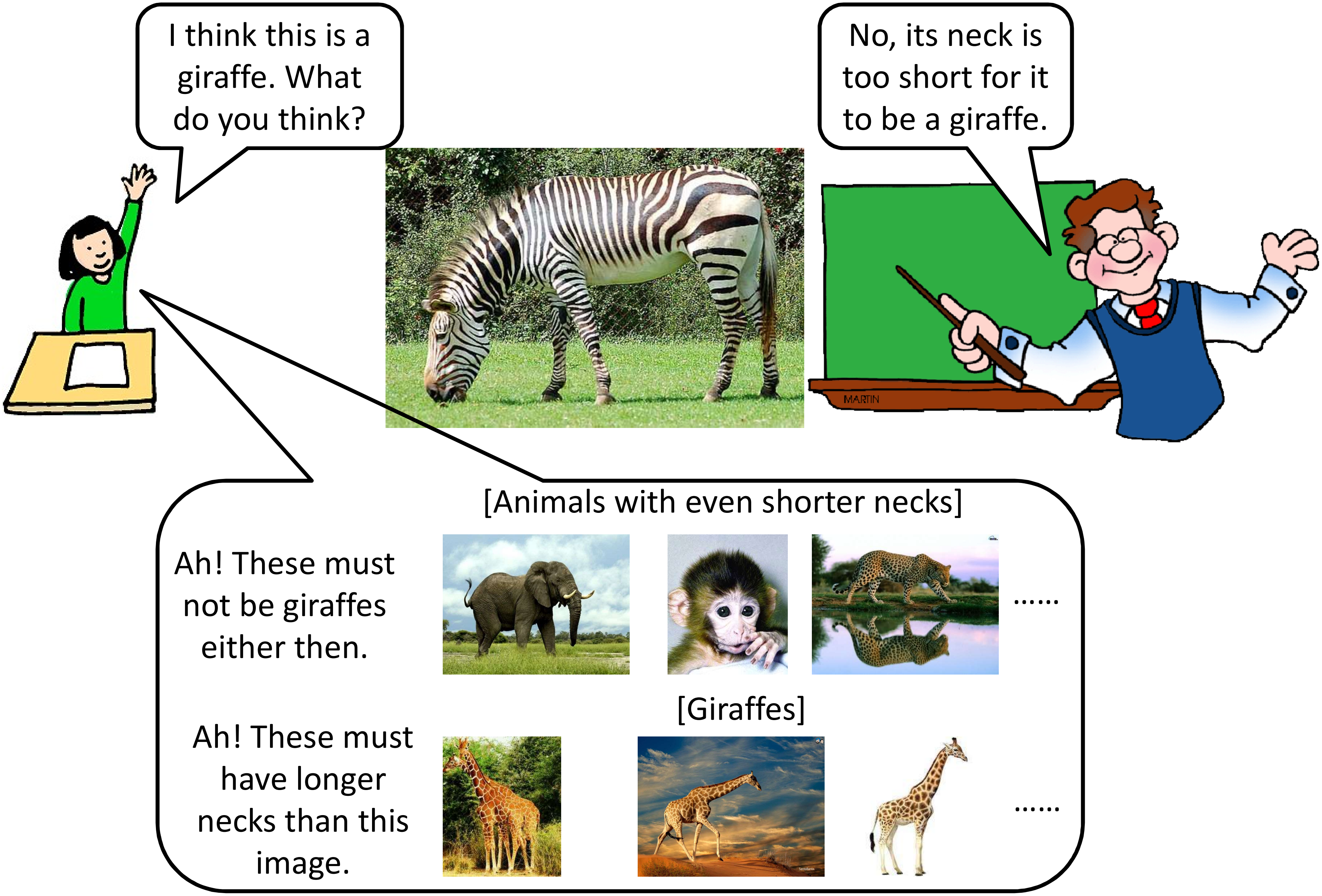
Attributes-based Feedback





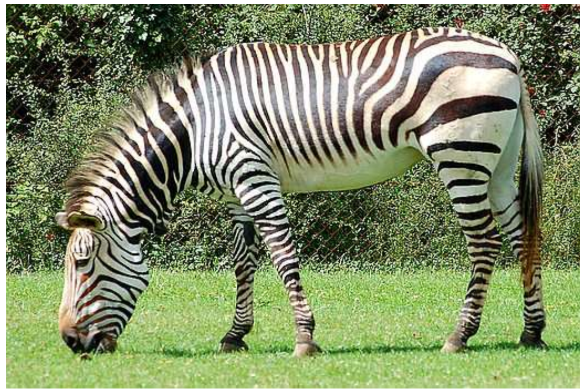
Slide credit: Devi Parikh

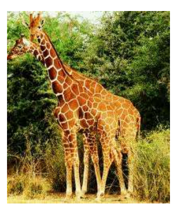
Biswas and Parikh, CVPR 2013



Slide credit: Devi Parikh

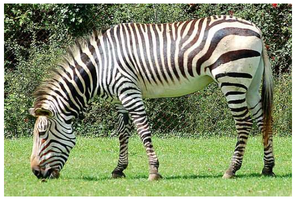
Biswas and Parikh, CVPR 2013

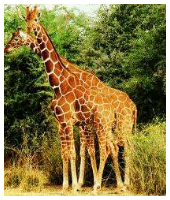


   ......

Slide credit: Devi Parikh

Biswas and Parikh, CVPR 2013

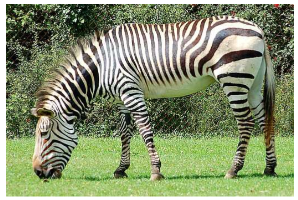


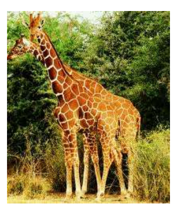


……

Slide credit: Devi Parikh

Biswas and Parikh, CVPR 2013



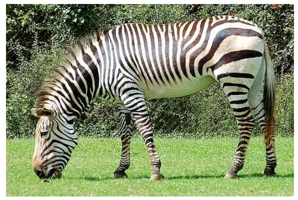



......

Slide credit: Devi Parikh

Biswas and Parikh, CVPR 2013

$\{$

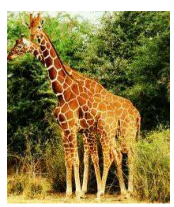
$($ $\succ$  $)$

$($ $\succ$  $)$

$($ $\succ$  $)$

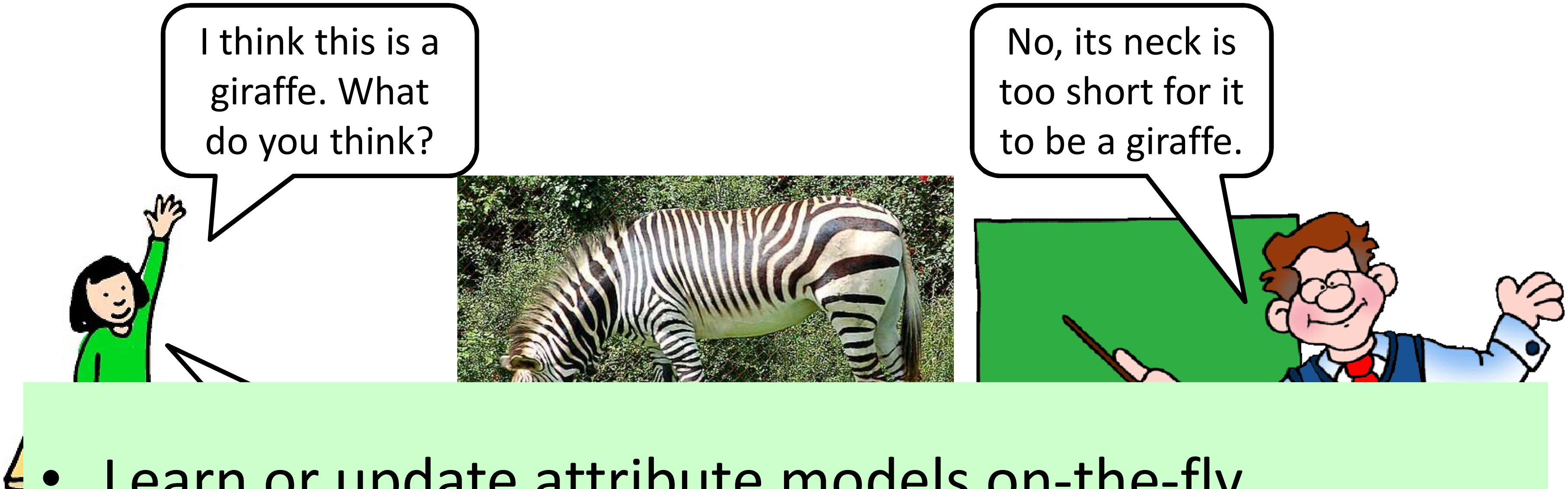
$\}$ $\bigcup O_m$

Update long-neck relative attribute

   ……

Slide credit: Devi Parikh

Biswas and Parikh, CVPR 2013



- Learn or update attribute models on-the-fly
- Start with an unlabeled pool of image and learn categories and attributes from scratch
- Actively select image for this form of feedback

Slide credit: Devi Parikh

Questions?