

Detecting Rare Events in Cognitive Systems

Misha Pavel, Daphna Weinshall,,
Hynek Hermansky, Holly Jimison and partners

Outline

- DIRAC
- Background: Novelty is old
- Oddballs vs. rare
- Categorization framework
- Rare definition
- Role of utilities
- Examples

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

DIRAC Preview: Problem Statement

- Detect and interpret (respond appropriately) to “rare” events
- Initial definition of rare:

  Clear but unexpected inputs in a given context
- Humans can do this usually very well, but machines fail
- Approach:
  - Detection based on comparison of inputs to expectations
  - Recognition based on fusion of multimodal inputs

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

DIRAC Circle of Partners

Leibnitz Inst Neurobiology



KATHOLIEKE UNIVERSITEIT LEUVEN



HUJI

CTU



CARL VON OSSIETZKY universität OLDENBURG

DIRAC
Detection and Identification of Rare Audiovisual Cues

BRNO UNIVERSITY OF TECHNOLOGY
FIT FACULTY OF INFORMATION TECHNOLOGY

OREGON HEALTH & SCIENCE UNIVERSITY

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITY OF MARYLAND
Neural Systems Laboratory

Fraunhofer FhG Gesellschaft

idiap RESEARCH INSTITUTE

Workpackages

- WP1: Signal Acquisition and Associated Processing
- WP2: Auditory Representation
- WP3: Visual Representations
- WP4: Learning and Categorization
- WP5: Multi-sensory Information Fusion
- WP6: Integration and Applications
- WP7: Training and Education
- WP8: Management, Planning and Dissemination

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Intelligence: Response to the Unknown

- Ubiquity of the problem: Response to the unexpected (fight or flight)
- Long history in philosophy and science from Aristotle to modern philosophers

- Psychology and Cognitive Science
  - Cognitive functions and intelligence
  - Discrimination
  - Categorization
  - Generalization
- Informatics Examples
  - Information Theory: Quantification of surprise
  - Statistical pattern recognition, classification
  - Machine Learning
  - Data Compression

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Engineering Problem

- Given: A classification system designed to respond optimally to examples in a training & validation sets
- System is confronted with a stimulus that is different from those in the training set sample distribution
- Possible responses:
  - Fixed Response: Best fitting class
  - Adaptive Response: Modify the best fitting class – adapt the most probable category
  - Adaptive Response: Create a new category with a temporary label
  - Adaptive Response: Run ☺

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Basic Premise

- Observation: Humans and animals are usually good at responding appropriately* to the detection of "rare" stimuli and events

- Violations of this observation are striking

- Can we build robots with these capabilities?
- Can we refine neuroscience paradigms to determine the neural substrate of this capability?

* Optimally with respect to given utility function

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Determinants of Intelligent (Objective) Response

- Is there an explicable reason for the stimulus interpretation?
  - Noise
  - Distortion
  - Occlusion
  - Context
- Is the response to the rare event important?
  - Context
  - Task
  - Consequences

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

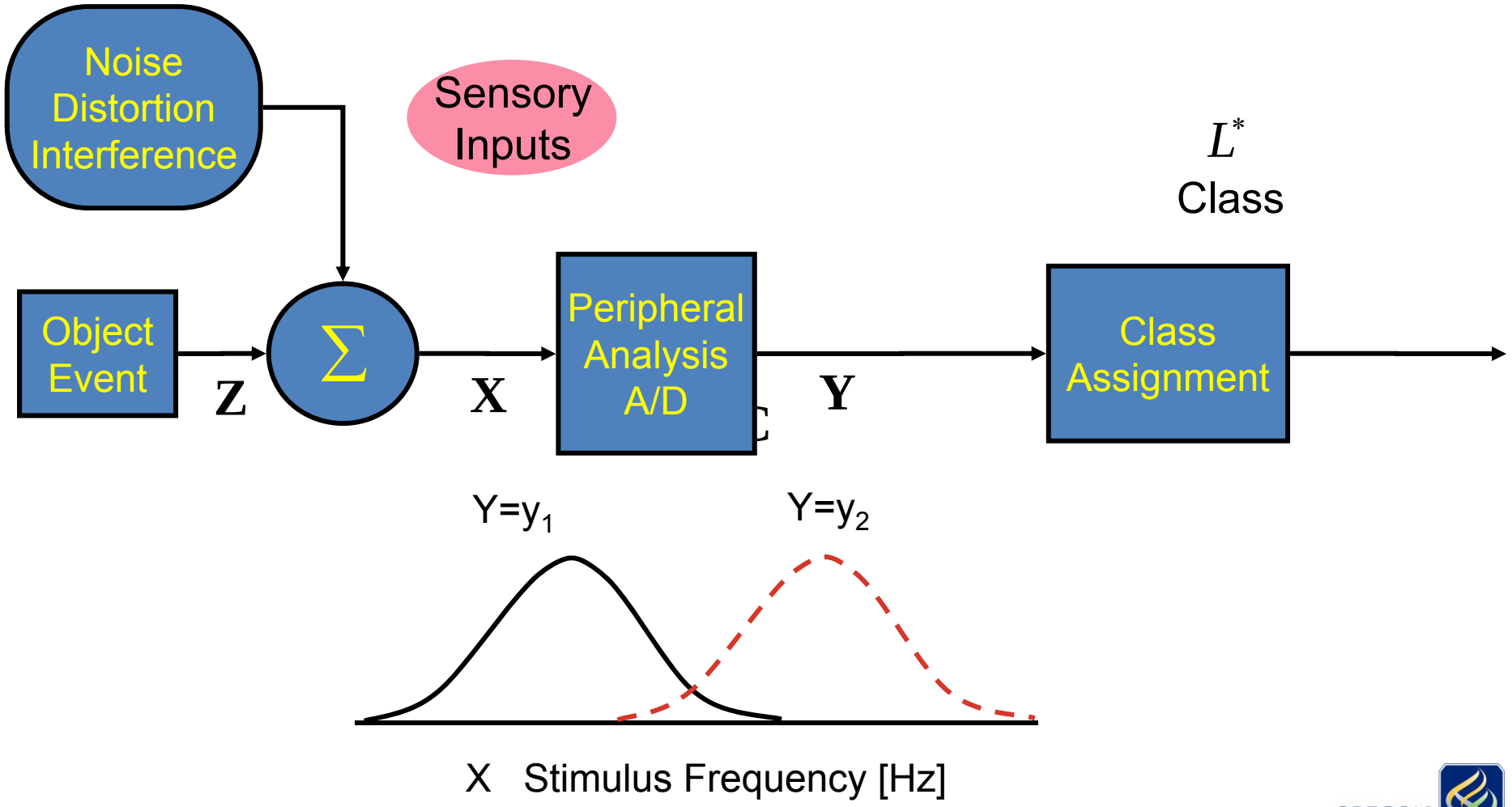
Psychology: Paradigms and Approaches

Paradigms used to study of humans and animal responses to novel stimuli

- Discrimination
  - Sensitivity of the sensory system
- Categorization
  - Search
  - Oddball detection
- Generalization

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

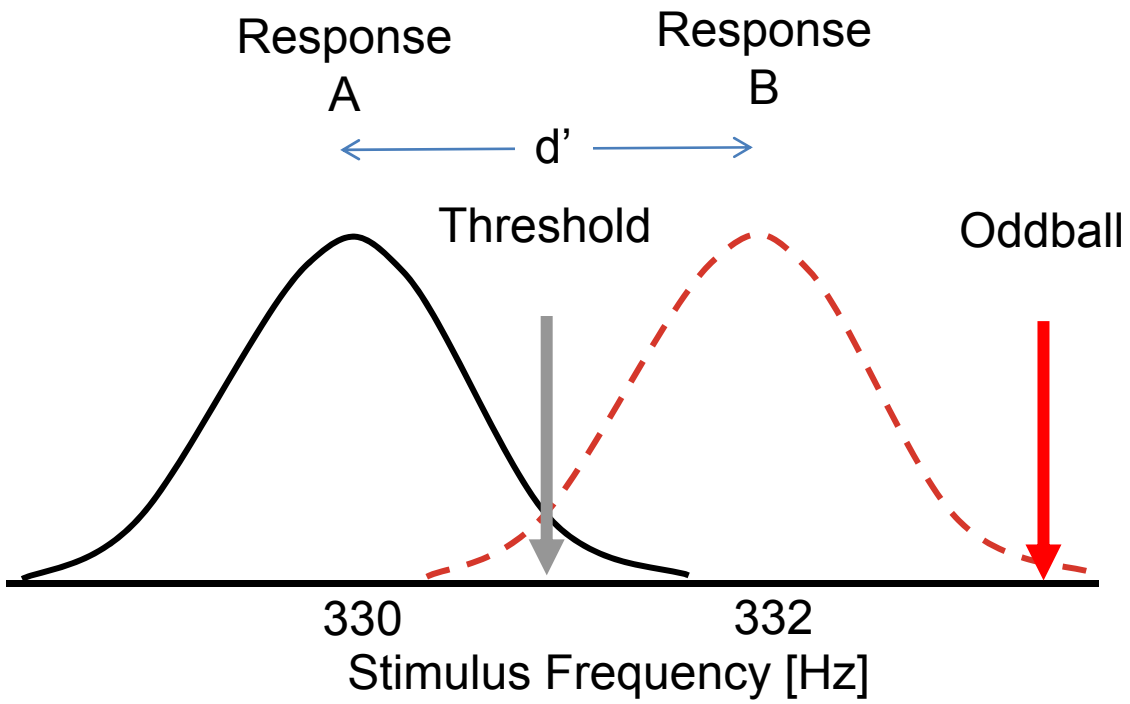
Framework for Discrimination and Categorization



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Perceptual Discrimination

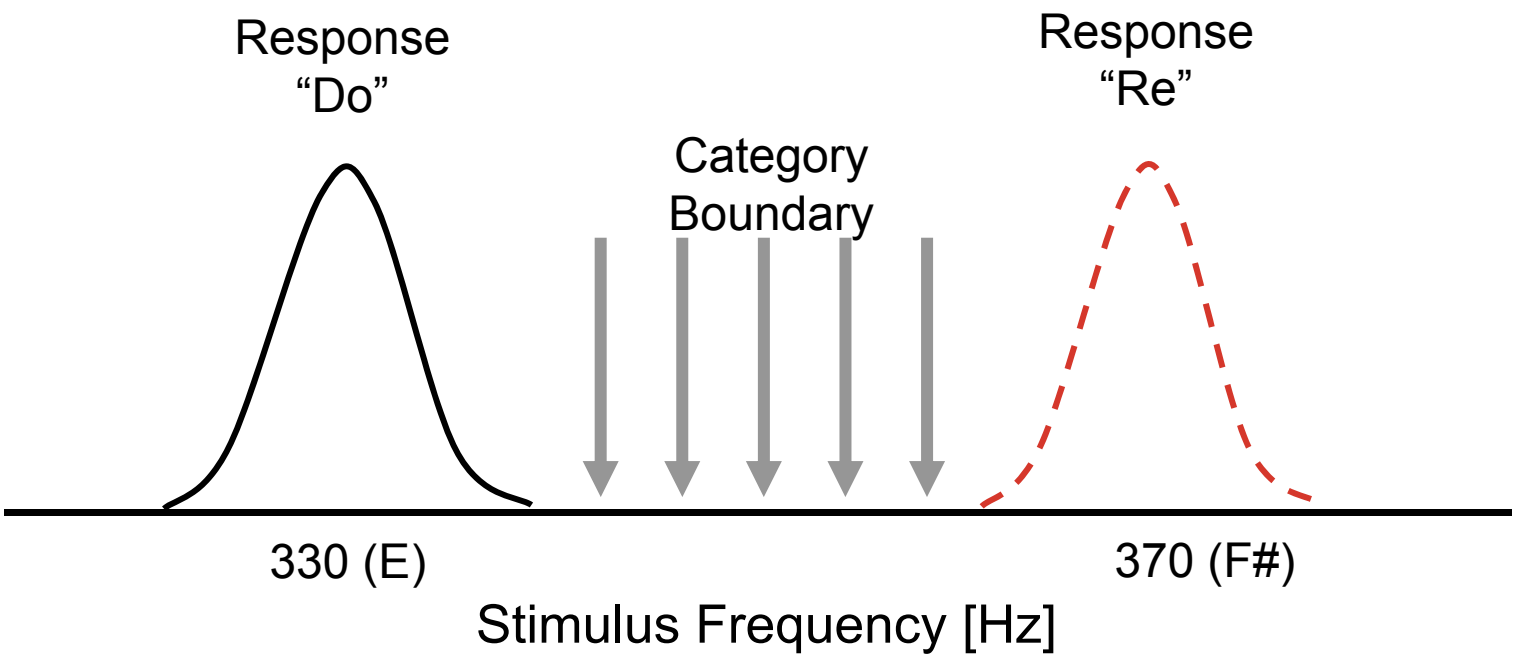
Example: Discrimination between two tones, f_1 and f_2



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Perceptual Categorization

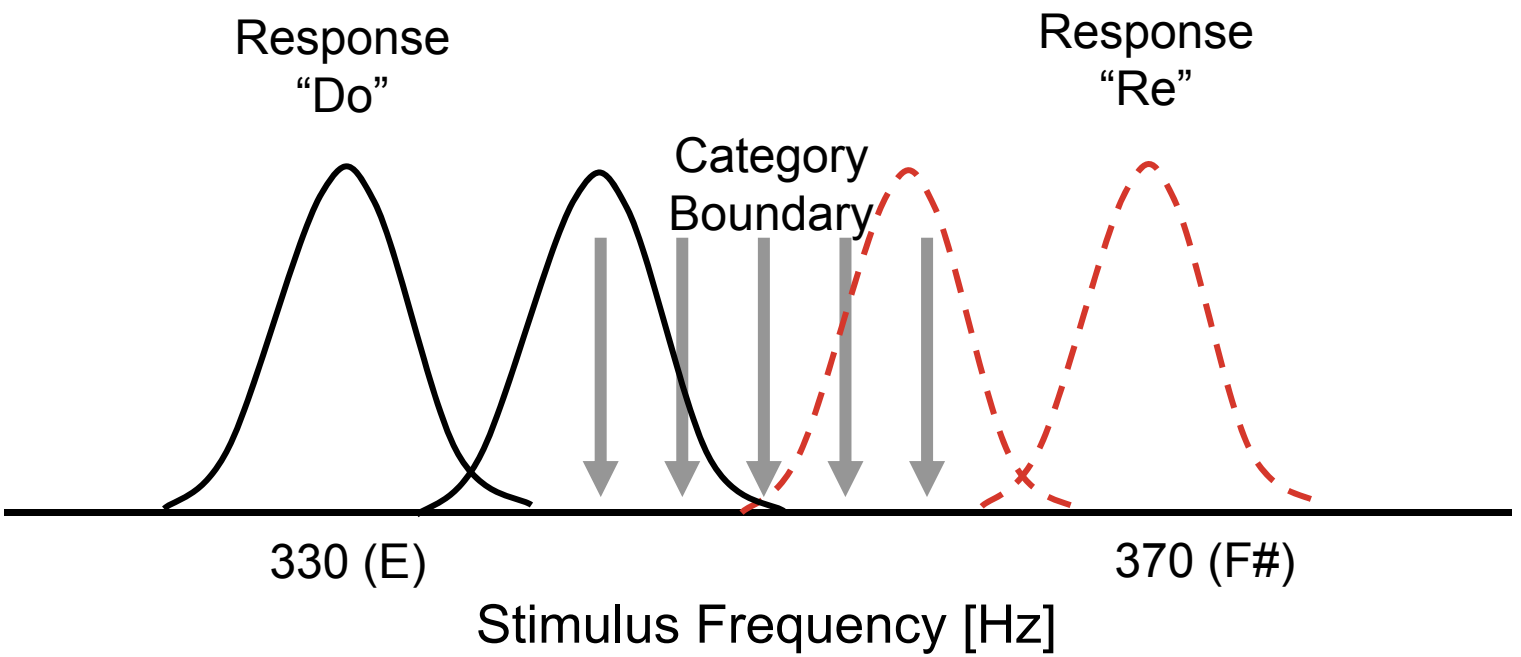
Example: Categorization of tones, f_1 and f_2



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Perceptual Categorization

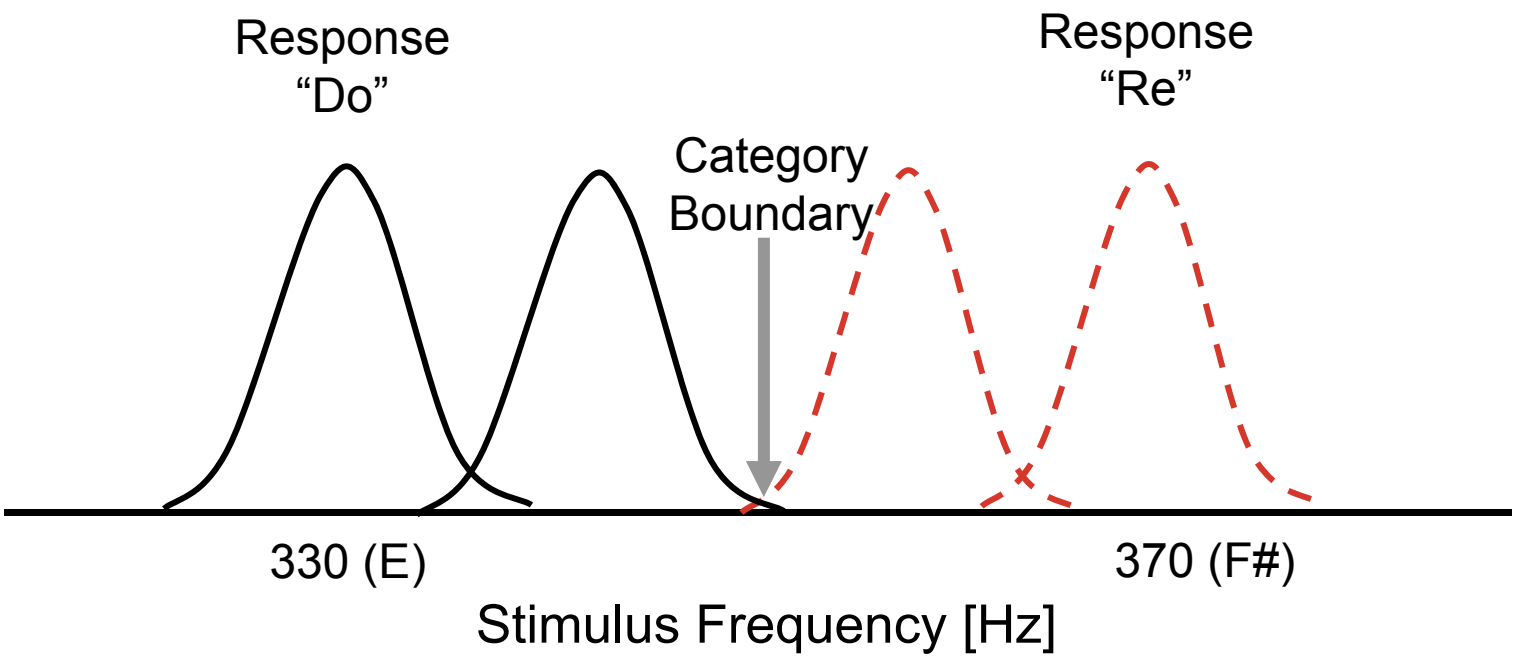
Example: Categorization of tones, f_1 and f_2



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

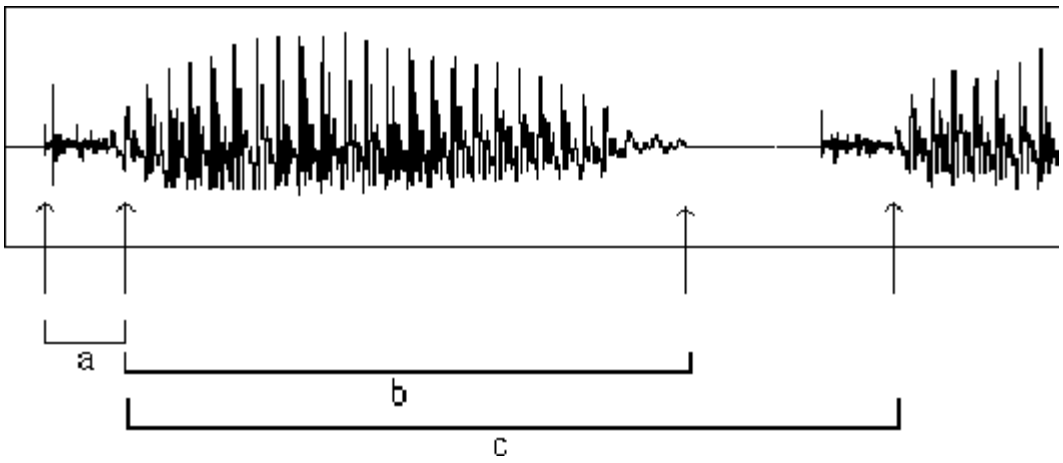
Perceptual Categorization

Example: Categorization of tones, f_1 and f_2



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example: Speech Sounds Discrimination

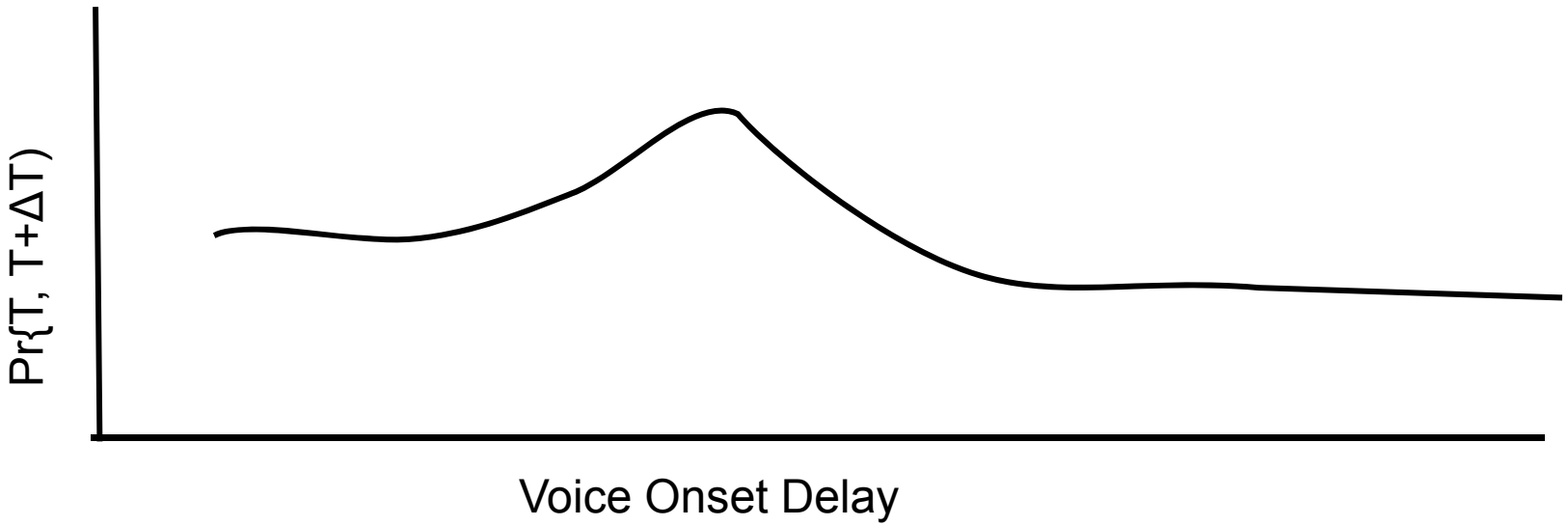


a. VOT (voice onset time), b. voicing duration, c. voice onset interval

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Categorical Perception?

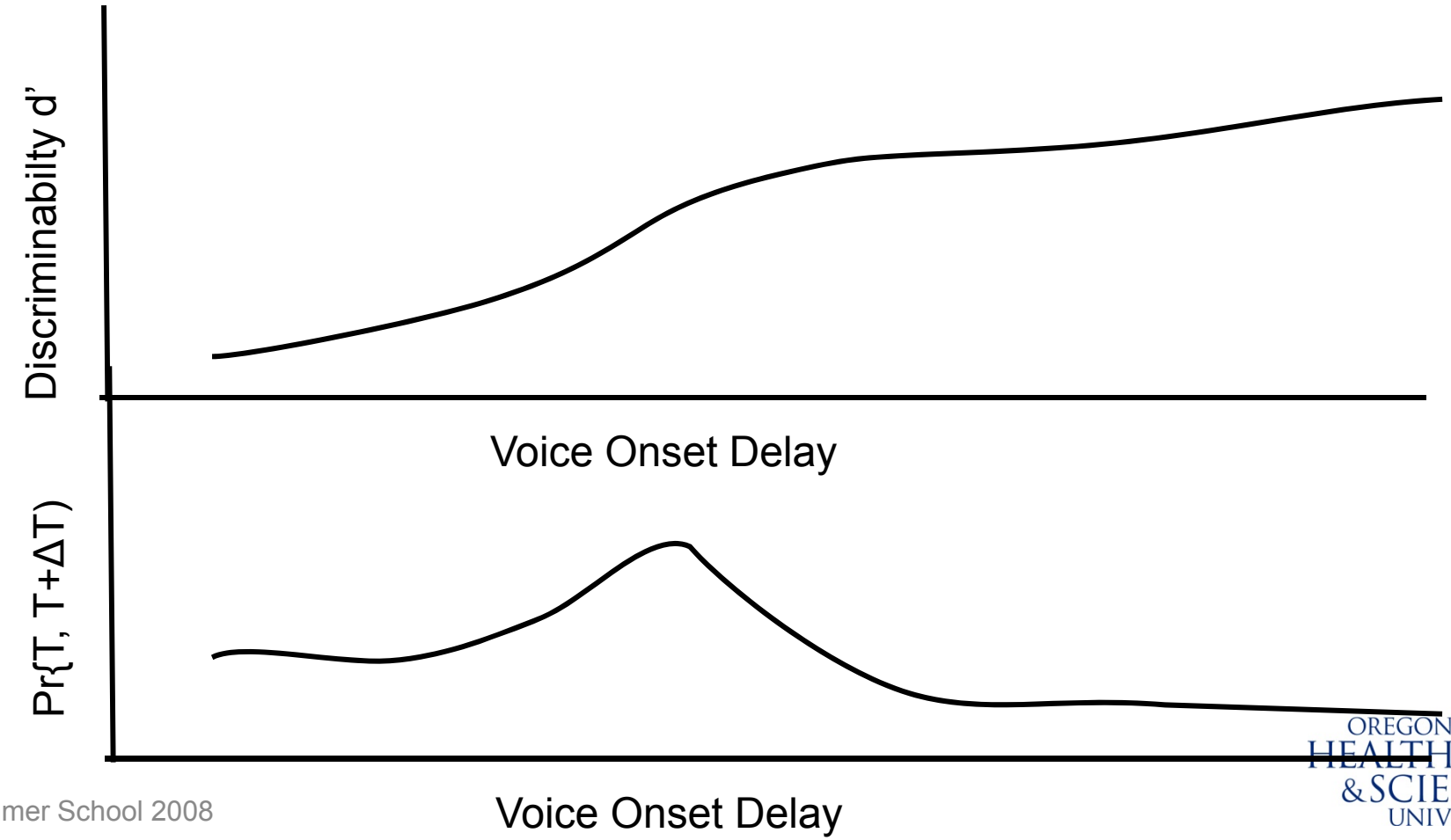
- Speech sound perception – discrimination
- Discrimination of Pr{T, T+ΔT) (b vs. p)



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Categorical Perception?

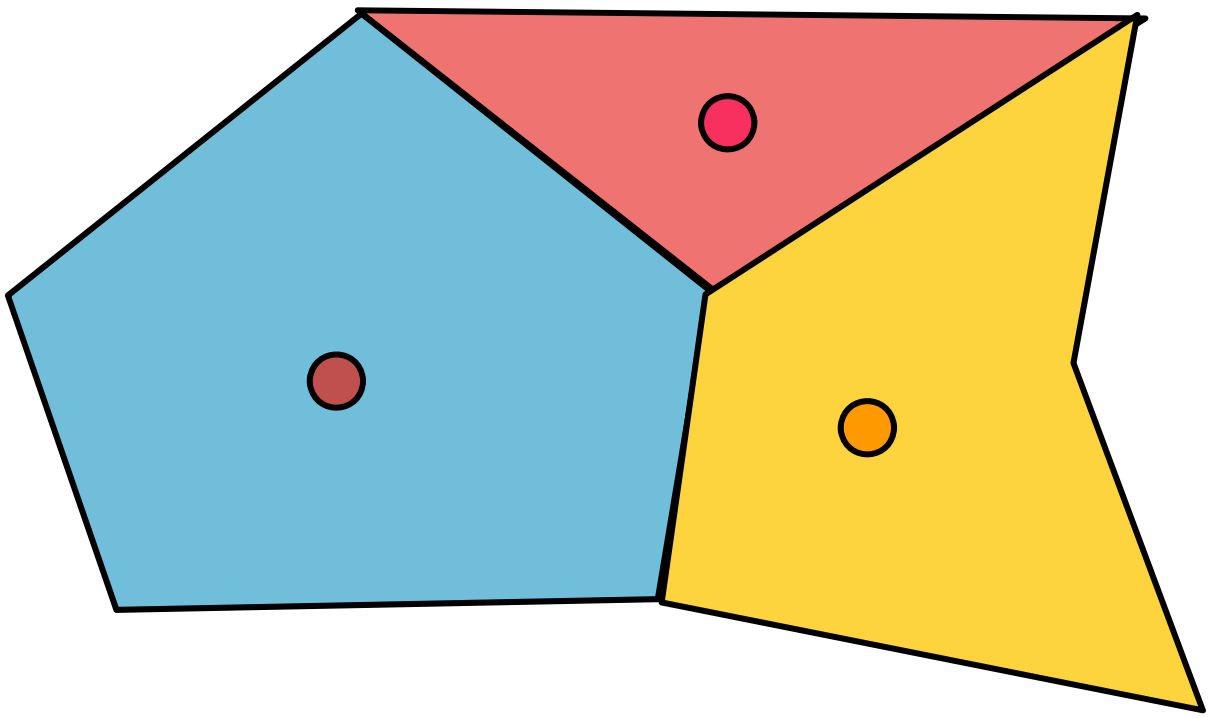
- Speech sound perception – discrimination
- Discrimination of Pr{T, T+ΔT)



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Multidimensional Categorization

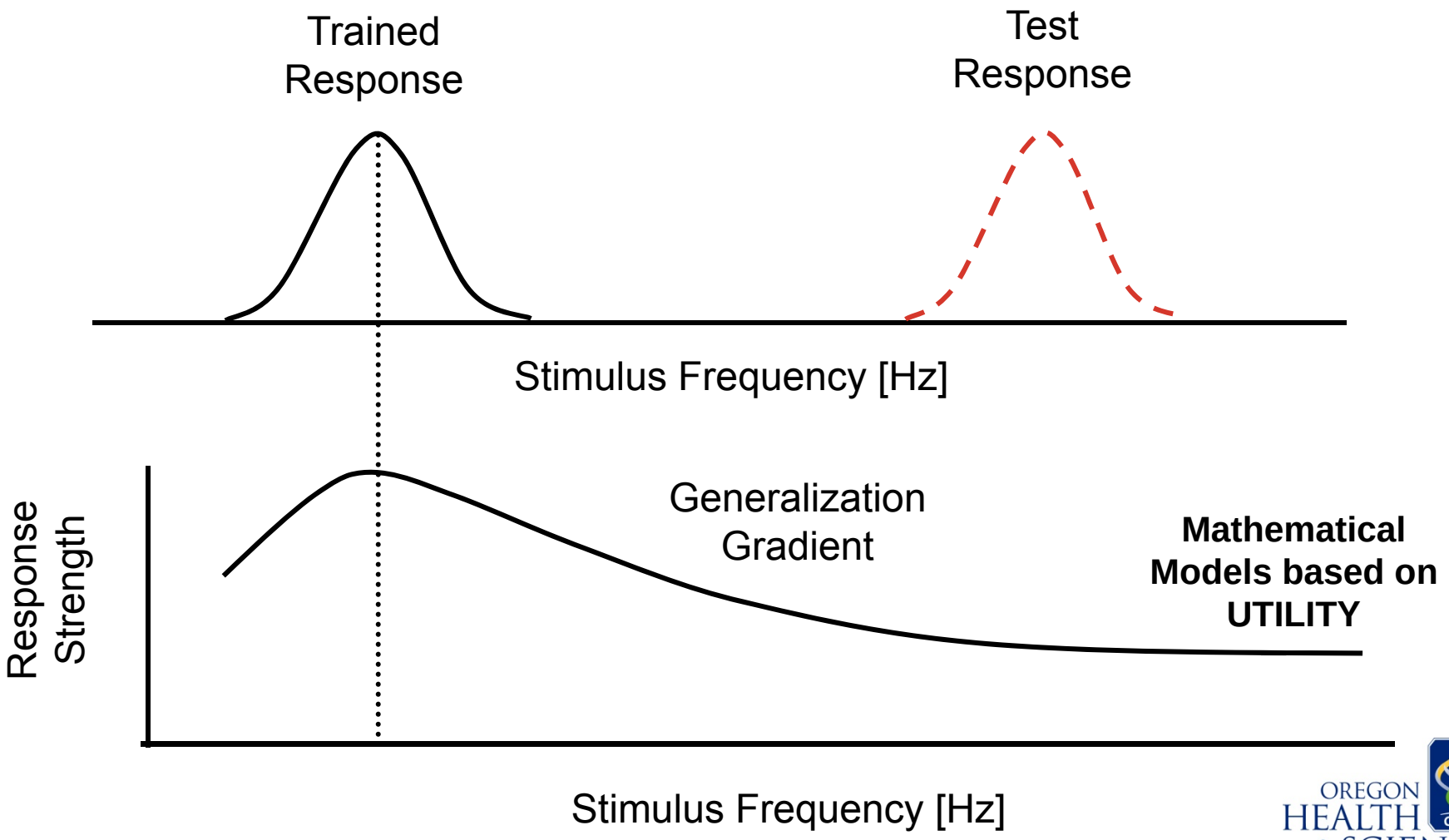
Example: Categorization Regions



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

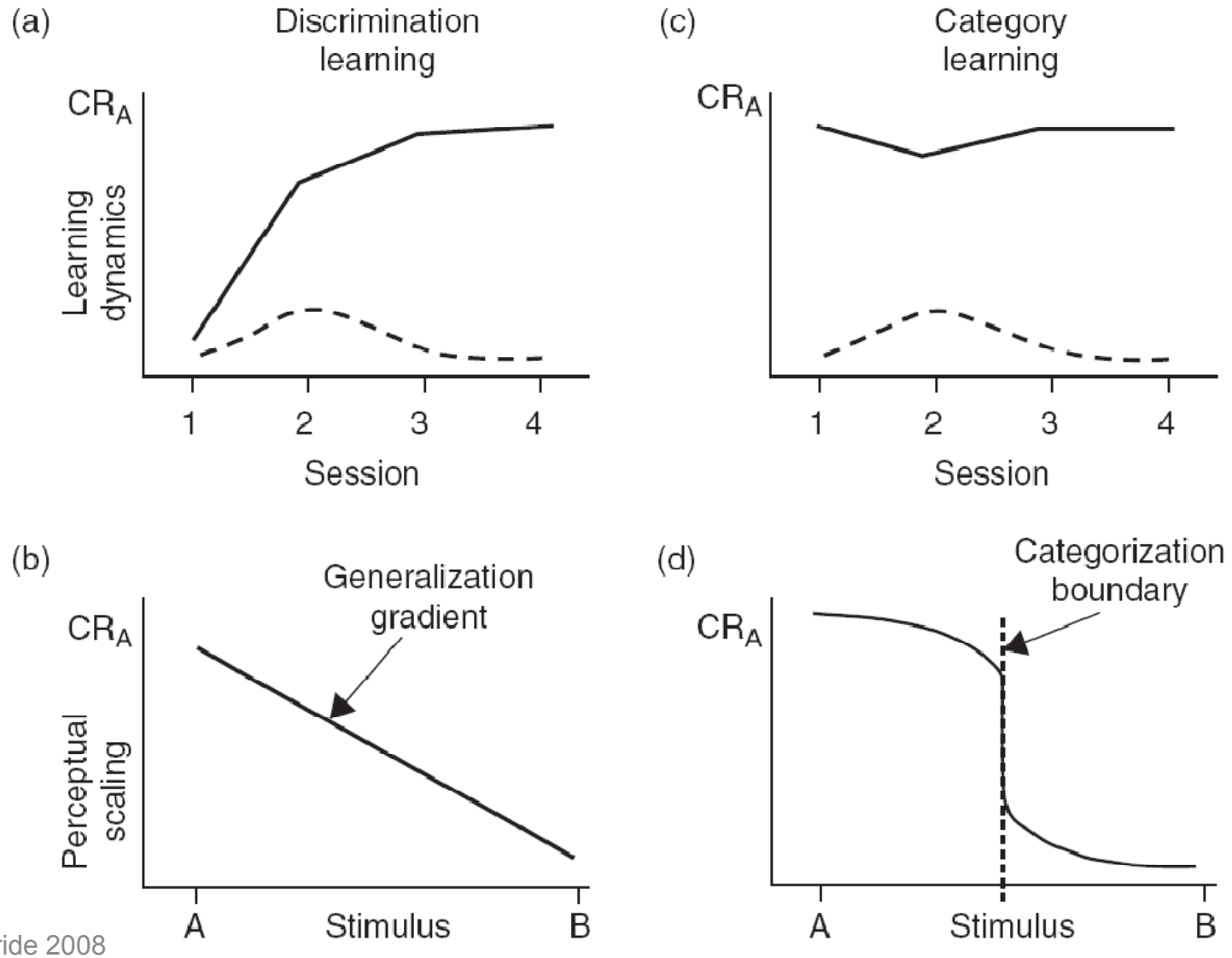
Psychological Generalization

Example: Salivation response to tones



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Generalization vs. Categorization

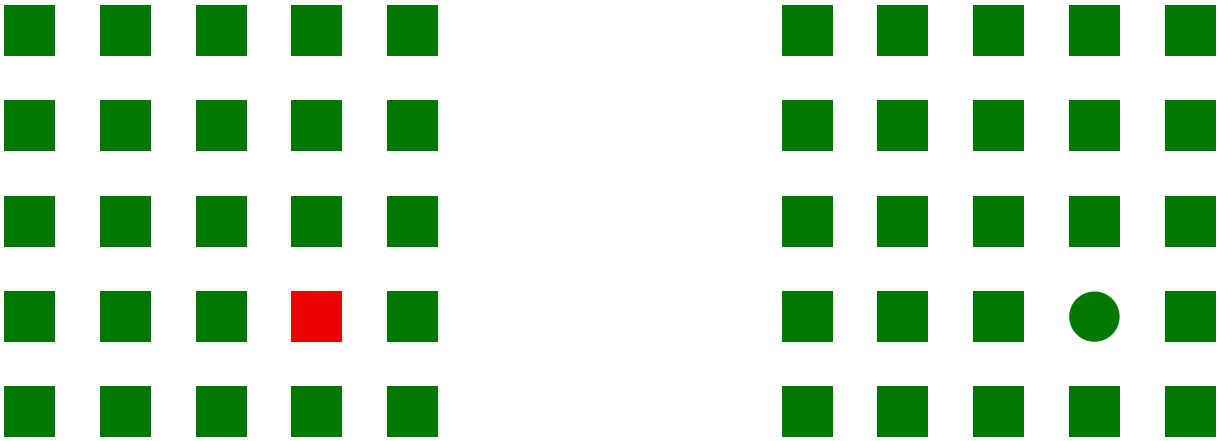


Ohl, 2008

OREGON HEALTH SCIENCE UNIVERSITY OHSU

Classical Definition of "Novel" Stimuli

Spatial Oddball Detection: Color or Shape



Temporal Oddball Detection: Color or Shape



Bottom-up, outlier detection

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Intuitive Notion of Rare, Unexpected Events

- Low **class** posterior probability can be caused by
  - Low prior probability
  - Uncertain, ambiguous measurement
  - Unexpected combination of observations
  - New class – to be added?
  - Low class prior probability in context
- Most current systems' response to Low probability stimuli
  - System finds the response with maximum a posteriori probability (MAP)
  - The output is the MAP response
  - System may provide confidence metrics
- Can system recognize its ignorance?
  Feature – class incongruence + high importance (utility)

Summary of Intuitive Description

- Low posterior probability due to conflicts among different interpretations of the same object or concept
- More generic interpretation has high posterior while the less generic has low posterior probability
- To make this intuition precise we need a formal structure

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Notation

- Observations $\mathbf{X} = \{X_1, X_2, ..., X_n\}$
- Features $\mathbf{Y} = \{Y_1, Y_2, ..., Y_m\}$
- Classes/Labels $\mathbf{L} = \{l_1, l_2, ..., l_k\}$
- Prior Class Probability $P\{L \mid C\}$
- Set of Utilities: $U = \{u_{00}, u_{01}, u_{10}.u_{11}\}$
- Context $C, \quad P\{C\}$
- Probability of new category $P\{L_{n+1} \mid C\}$

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Classification in Context

- Maximum a Posteriori --$P\{L \mid \mathbf{Y}, C\} = P\{\mathbf{Y} \mid L, C\} \dfrac{P\{L \mid C\}}{P\{\mathbf{Y} \mid C\}}$

  - Mode of posterior probability distribution

$$L^* = \arg\max_{\forall l} \left\{ P\{L \mid \mathbf{Y}, C\} \right\}$$

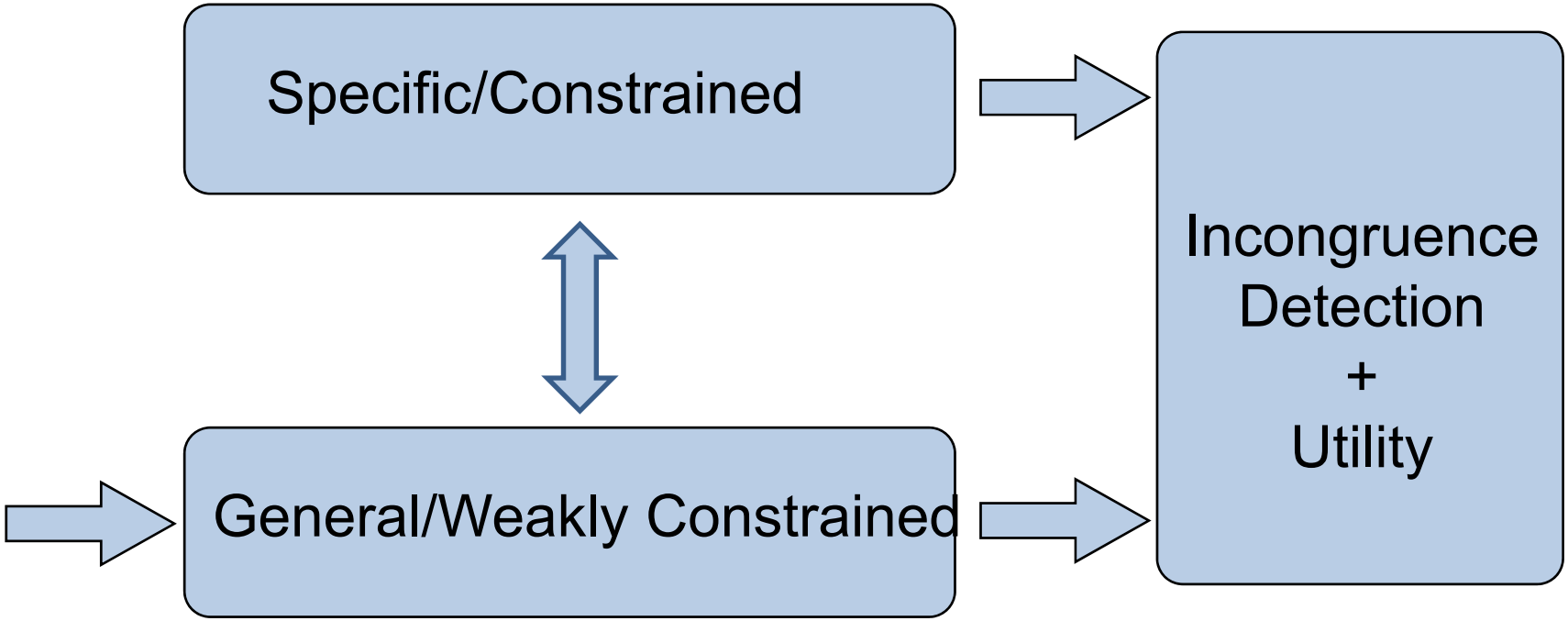
- Maximum Expected Utility:

$$L^* = \arg\max_{\forall L} \left\{ \sum_K u_{LK} P\{L \mid K, \mathbf{X}, C\} \right\}$$

Need a framework for classification

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

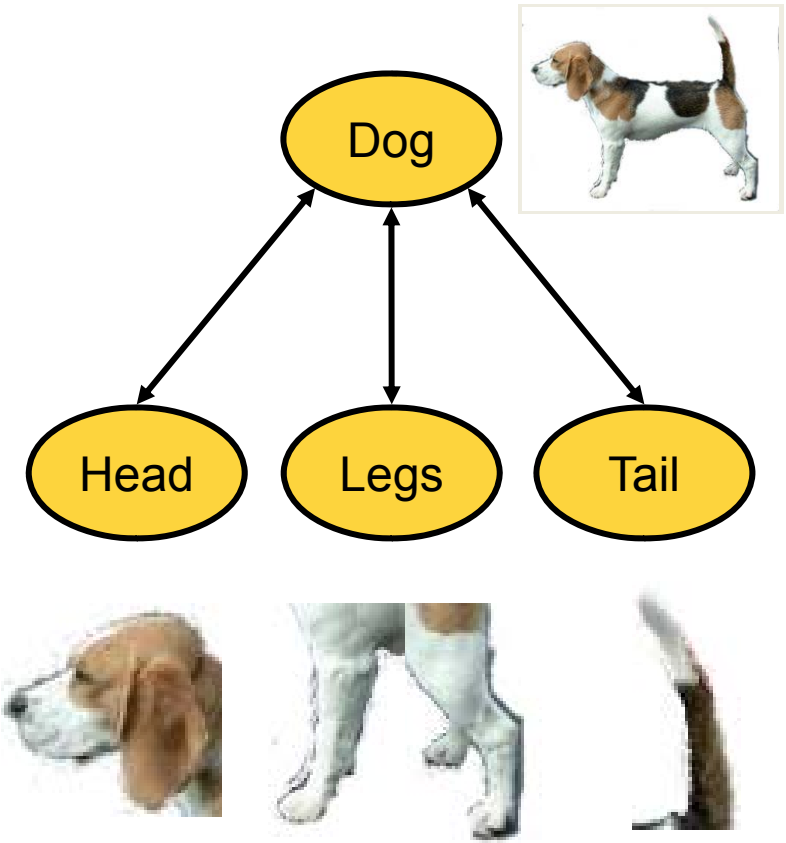
Framework: Detection of Conflict in Probabilities



Need a framework for the models and their relationships

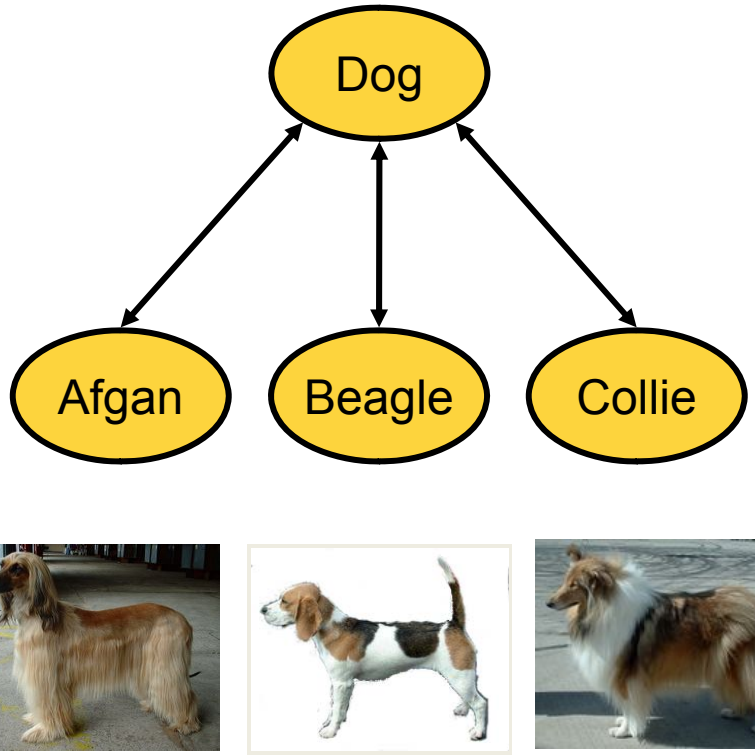
OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Part-Membership Hierarchy of Categories



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Class-Membership Categories

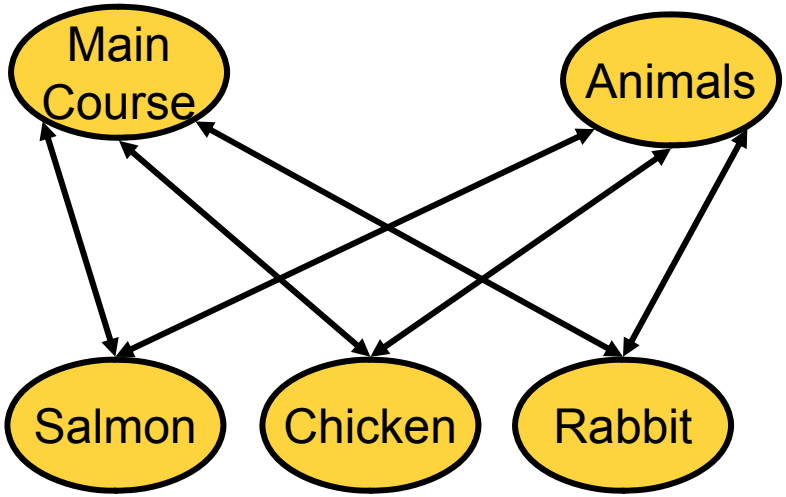


OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Shortcomings of a-priori Hierarchical Structures

- Strict hierarchy (tree-based representation) is violated
- Infinite number of levels – What level is appropriate?
  - Basic categories????
  - Well-defined hierarchy level???

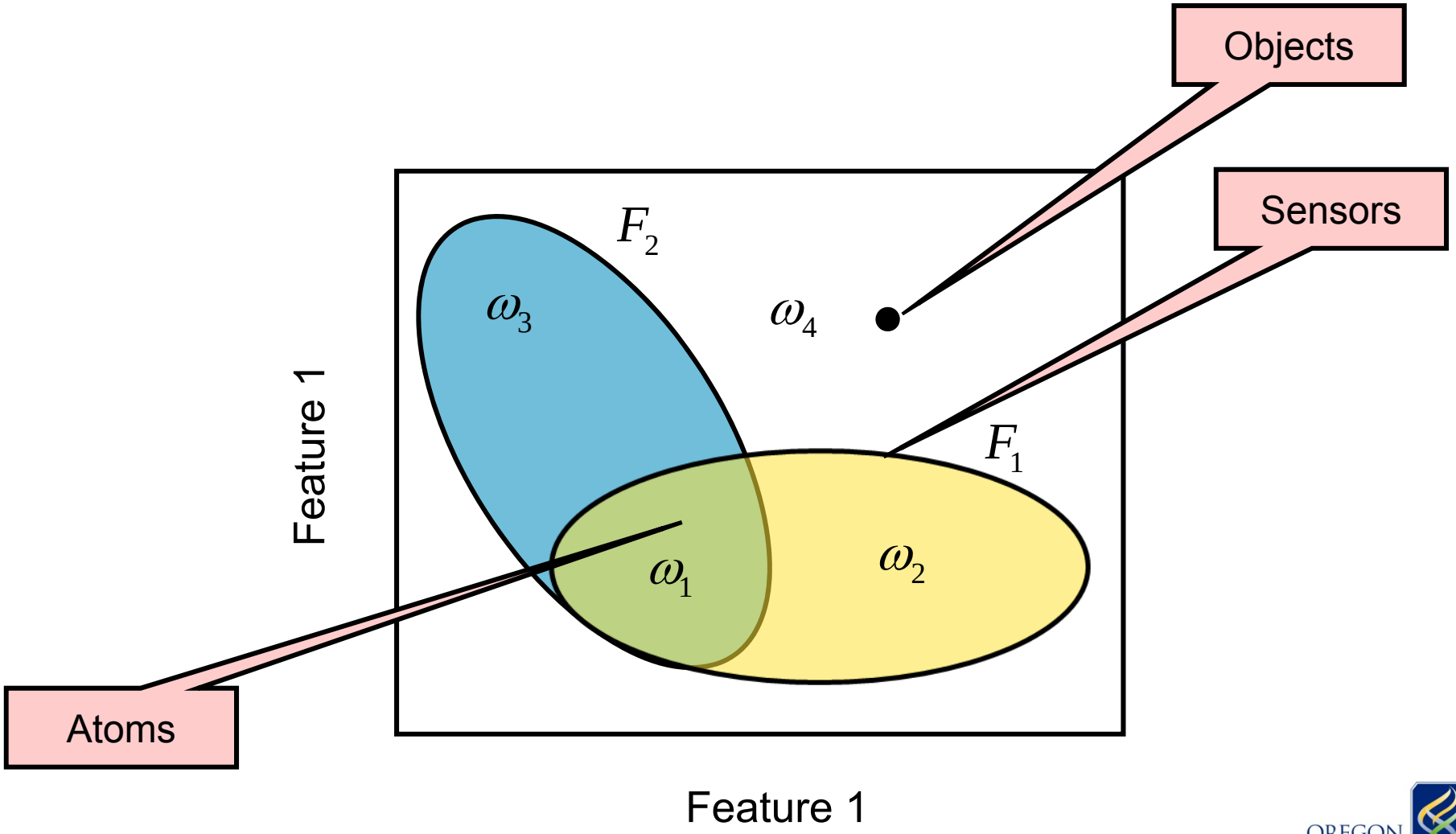
Depends on Context and Task



Can we develop a structure that captures the advantages hierarchy and overcomes
Incorporate context

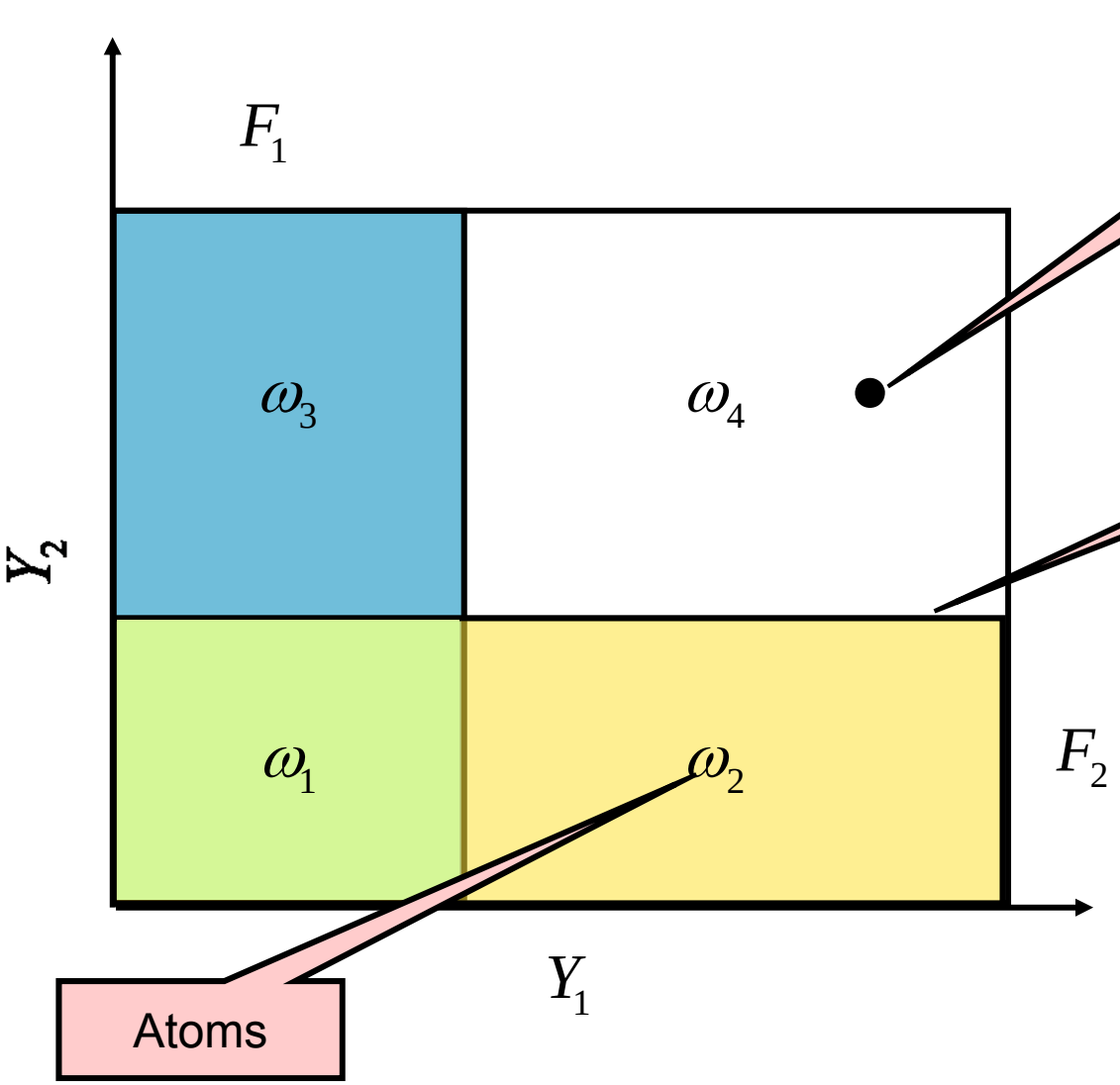
OREGON HEALTH & SCIENCE UNIVERSITY OHSU

An Alternative: Object – Feature Space

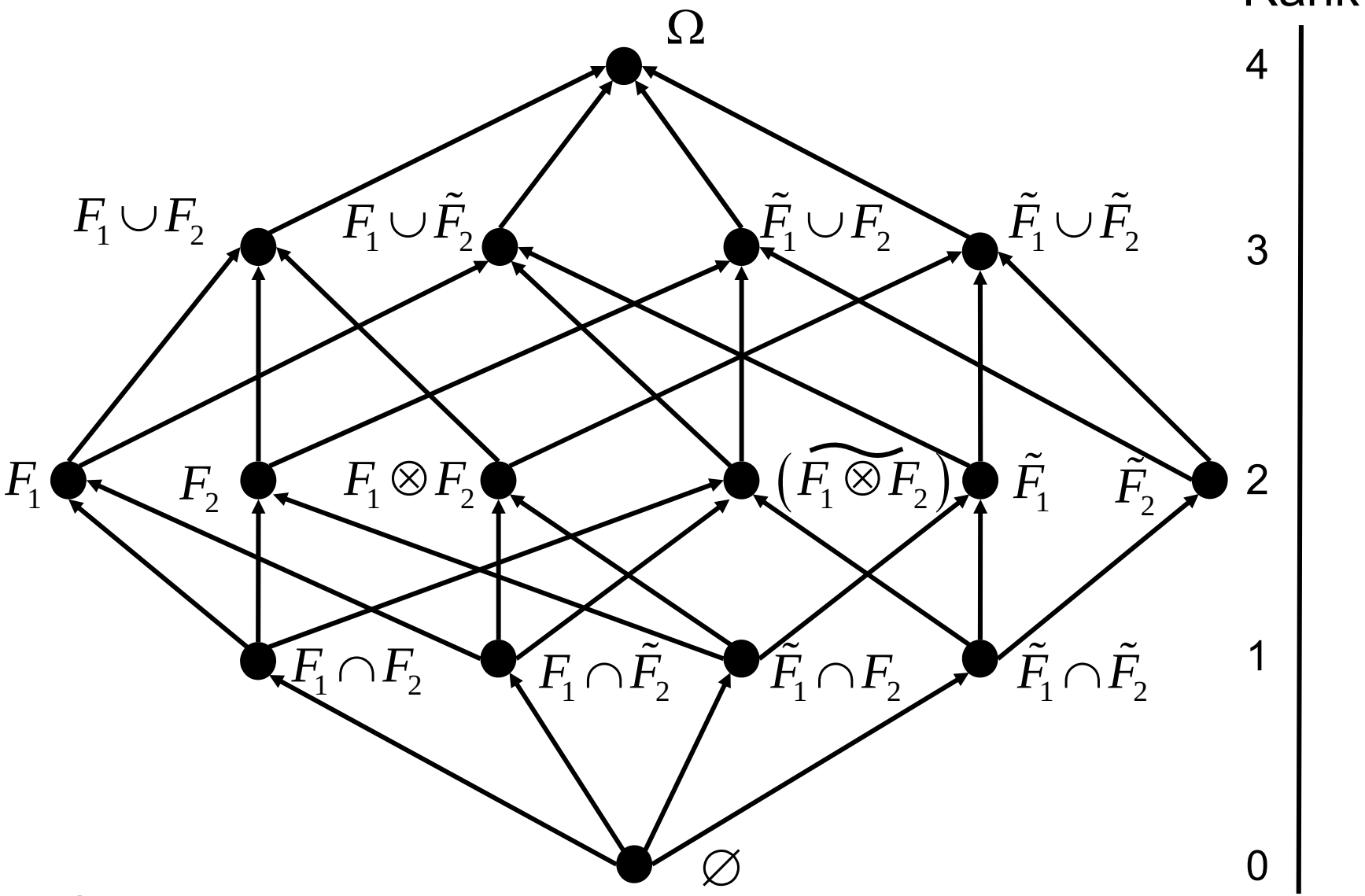


OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

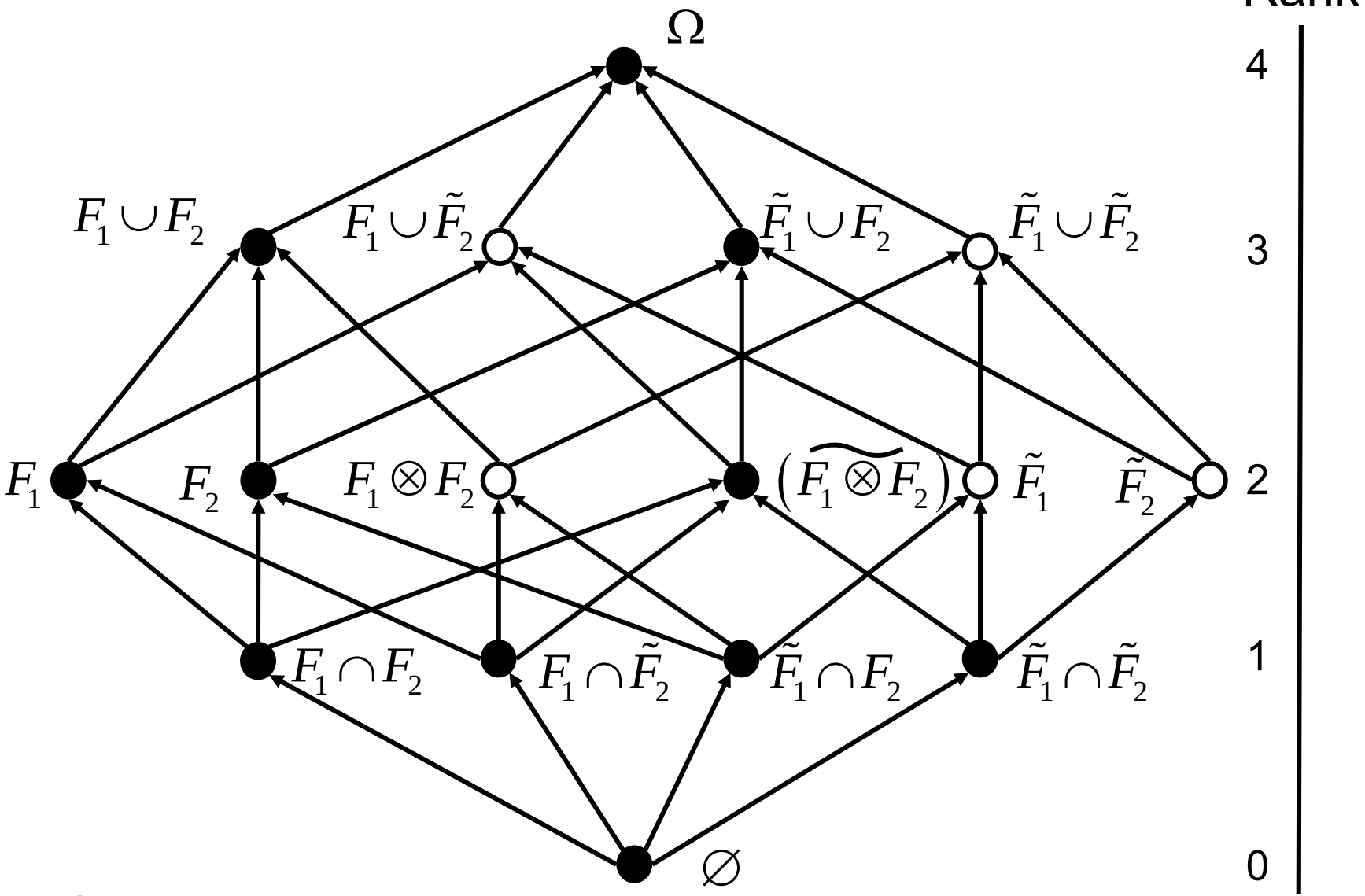
Object – Feature Space – Sensors as Predicates



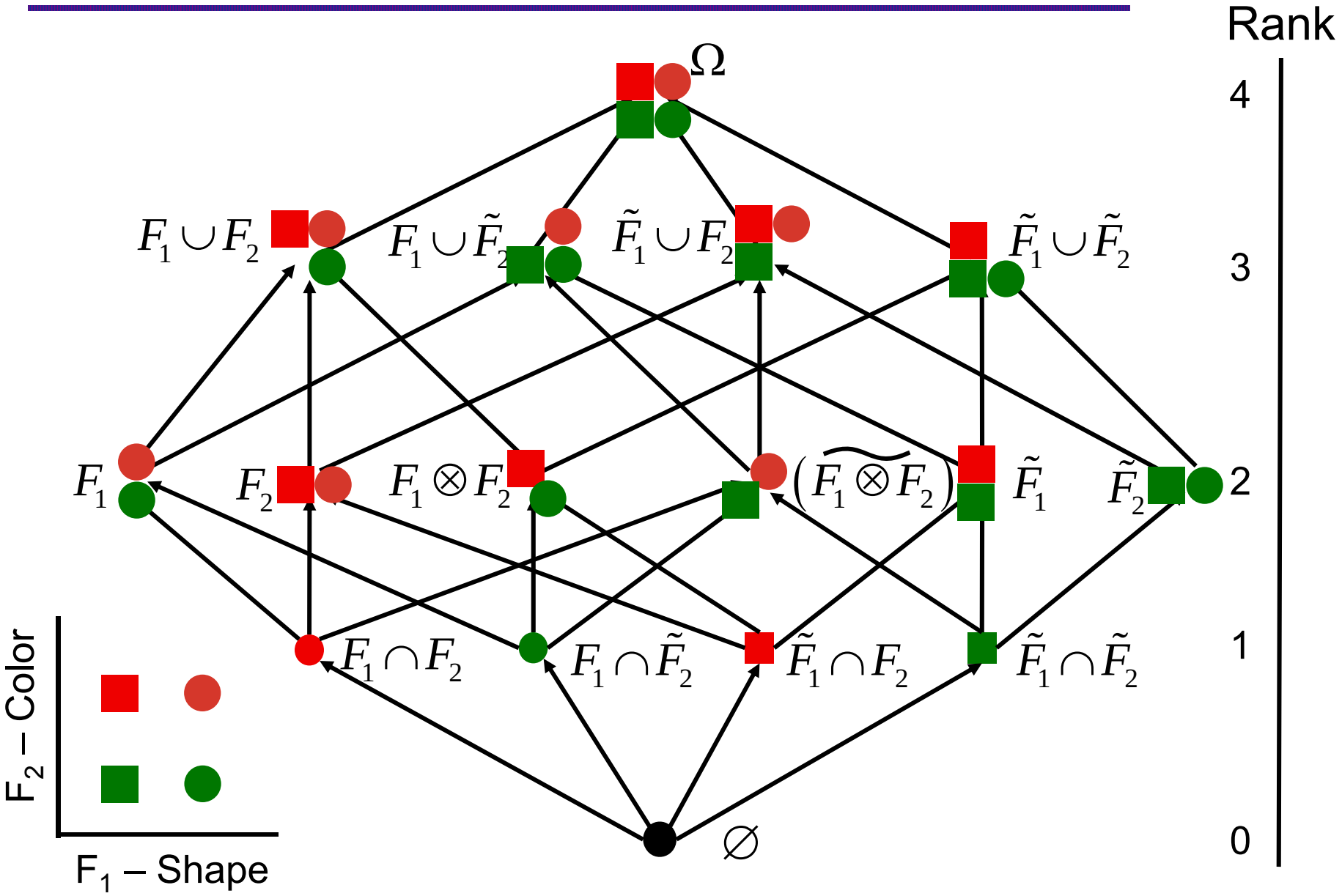
Complete Partial Order



Class-Specific Partial Order



Example: Partial Order in Visual Search



Object – Feature Space – Sensors as Predicates

Objects $y \in \mathbf{Y}$

Predicates $F_i : \mathbf{Y} \rightarrow \{\mathbf{0},\mathbf{1}\}$

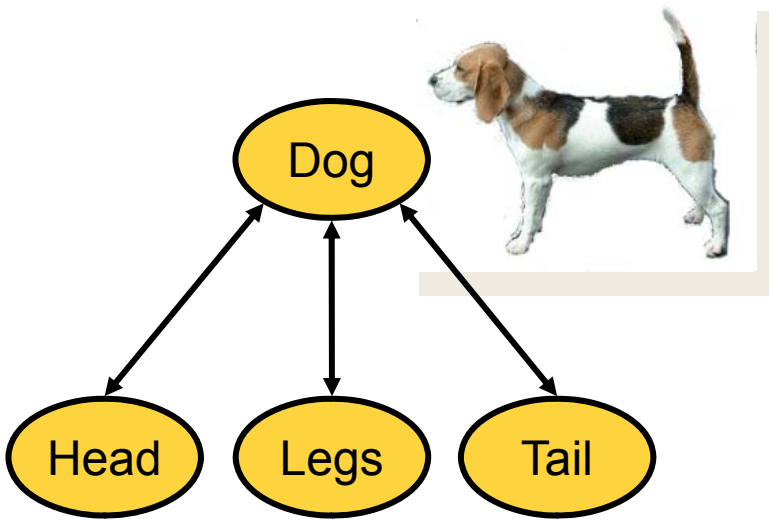
Partial Order $a \prec b \Leftrightarrow F_a \subseteq F_b$

Stronger
More Specific
Model

Weaker
More General
Model

OREGON HEALTH SCIENCE UNIVERSITY OHSU

Example: Part – Membership Category



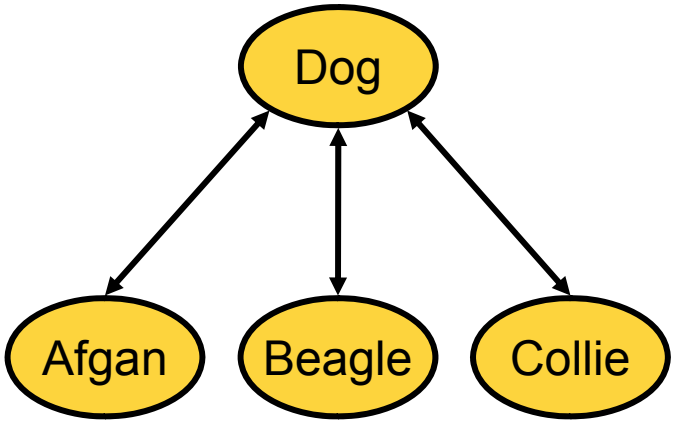
$$F_{dog} = F_{legs} \cap F_{head} \cap F_{tail}$$

$$F_{dog} \subset F_{legs} \Rightarrow dog \prec legs$$

Dog is more specific (stronger) model then parts

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example: Class – Membership



$$F_{dog} = F_{Afgan} \cup F_{Beagel} \cup F_{Collie}$$

$$F_{dog} \supset F_{legs} \Rightarrow dog \succ Afgan$$



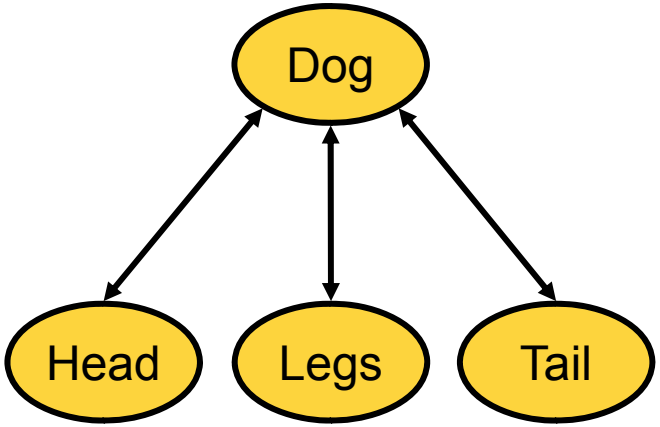




Dog is more general (weaker) model then the breeds

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Incongruent Part – Whole Category



Incongruent

$$F_{dog} = F_{legs} \cap F_{head} \cap F_{tail}$$

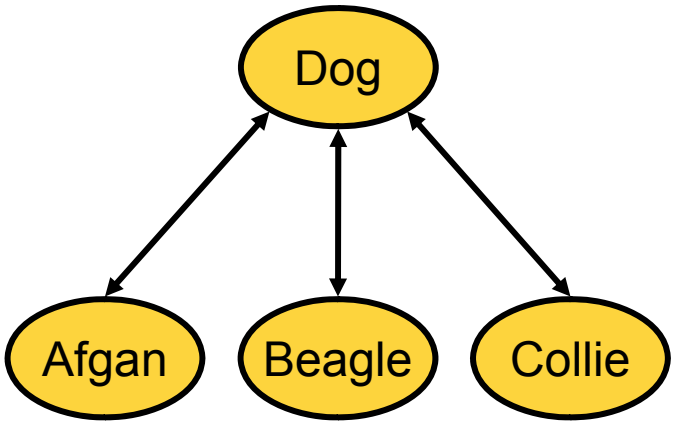
$$F_{dog} \subset F_{legs} \Rightarrow dog \prec legs$$

$$P^{s}_{dog} = \prod_{b \in A^{l}} P_b = P_{legs} P_{head} P_{tail}$$

$$\mathrm{P}^{s}_{dog}(\mathbf{X}) \gg \mathrm{P}_{dog}(\mathbf{X})$$

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Example: Class – Membership





Incongruent

$$F_{dog} = F_{Afgan} \cup F_{Beagel} \cup F_{Collie}$$

$$F_{dog} \supset F_{legs} \Rightarrow dog \succ Afgan$$

$$F_{dog} = F_{Afgan} \cup F_{Beagel} \cup F_{Collie}$$

$$\mathrm{P}_{dog}^{g} = \sum_{b \in A^{s}} P_{b} = P_{Afgan} + P_{Beagle} + P_{Collie}$$

$$\mathrm{P}_{dog}^{g}(\mathbf{X}) \ll \mathrm{P}_{dog}(\mathbf{X})$$

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

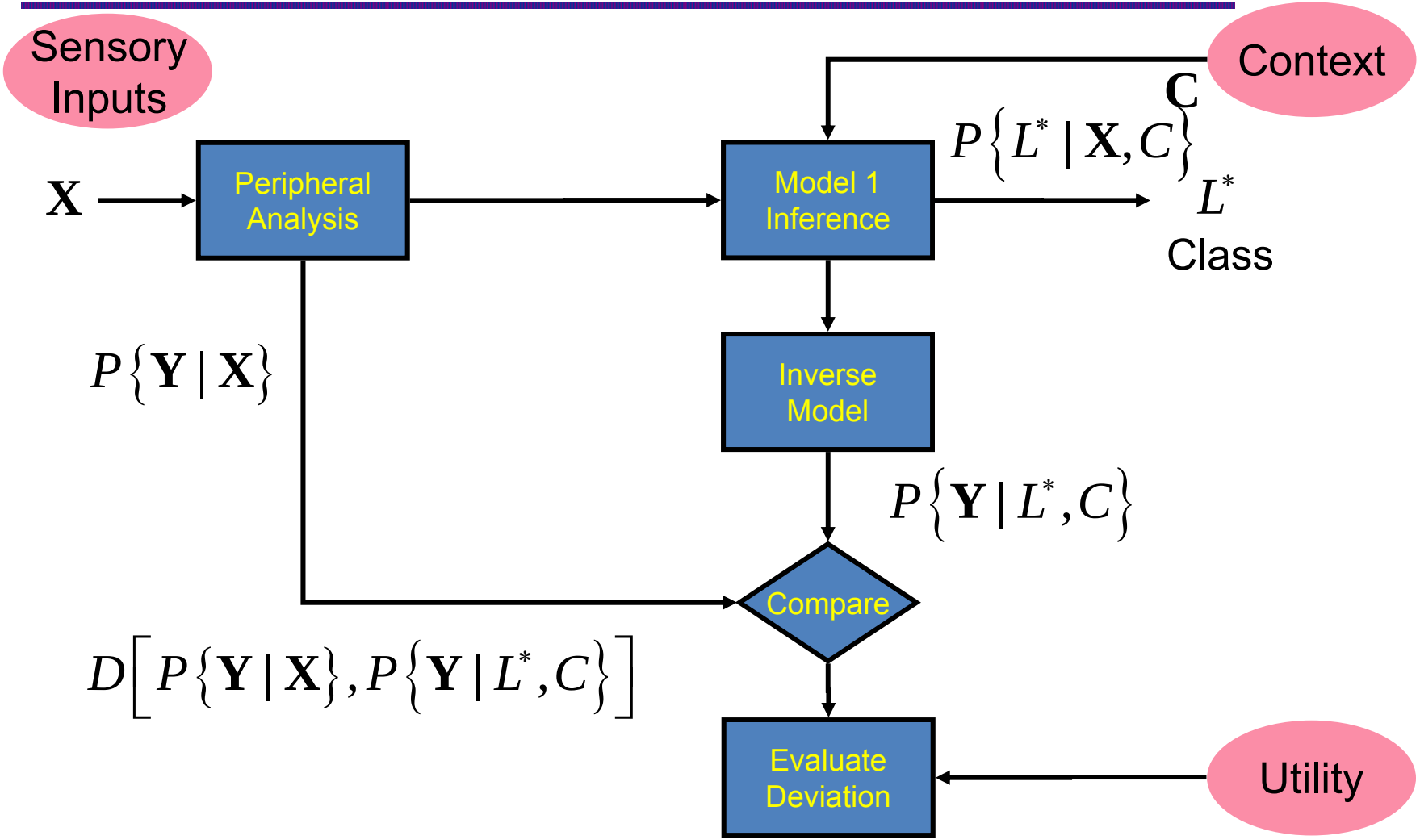
Rare – Incongruent Events

	Specific	General
"Noise" or oddball	Low	Low
Incongruent	Low	High
Incorrect Model	High	Low
Expected	High	High

$$D\left[\mathrm{P}_a^l(\mathbf{X}) \,|\, \mathrm{P}_a(\mathbf{X})\right] = \int \mathrm{P}_a^l(x) \log \frac{\mathrm{P}_a(x)}{\mathrm{P}_a^l(x)} dx$$

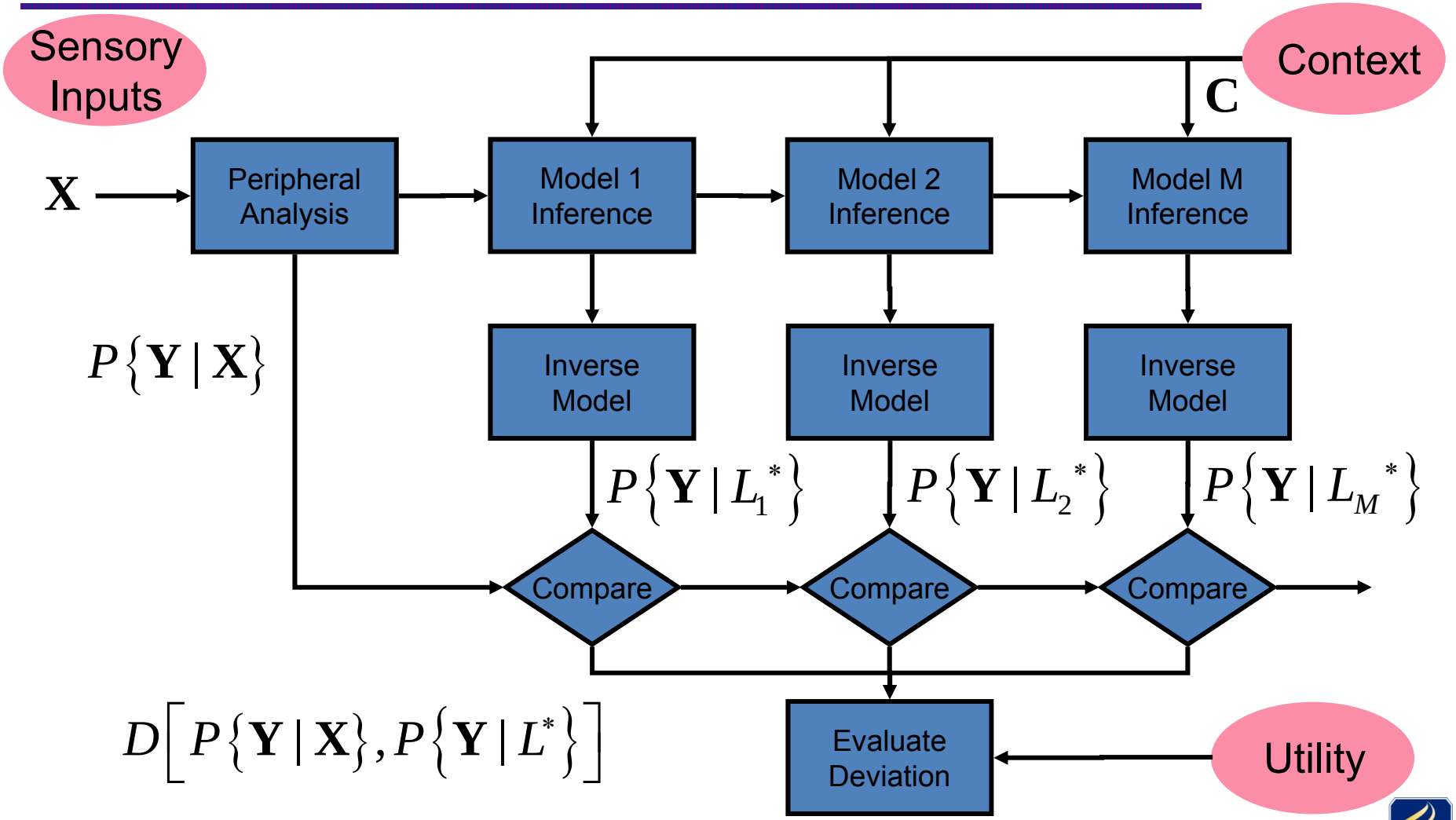
OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Algorithms for Detection of Rare Events



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Hierarchical Models of Rare Events



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Context

- Task
- Environmental setting
- Hierarchy of models
- Prior probabilities
- Utilities

Examples of Human Failures to Consider Utilities

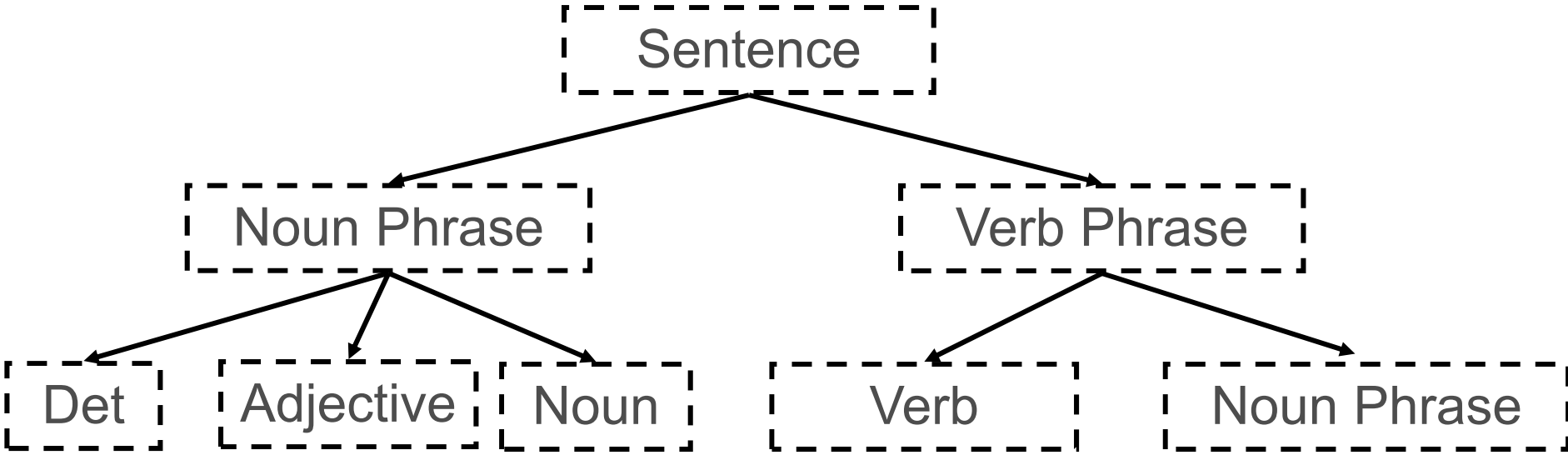
OREGON HEALTH & SCIENCE UNIVERSITY OHSU

How to Get Utility Estimates?

- Utility estimation from context and background
  - Linguistic analysis
  - Multimodal inputs
  - Contextual cues
  - Task objectives

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Example of Utility Assessment: Linguistic Analysis



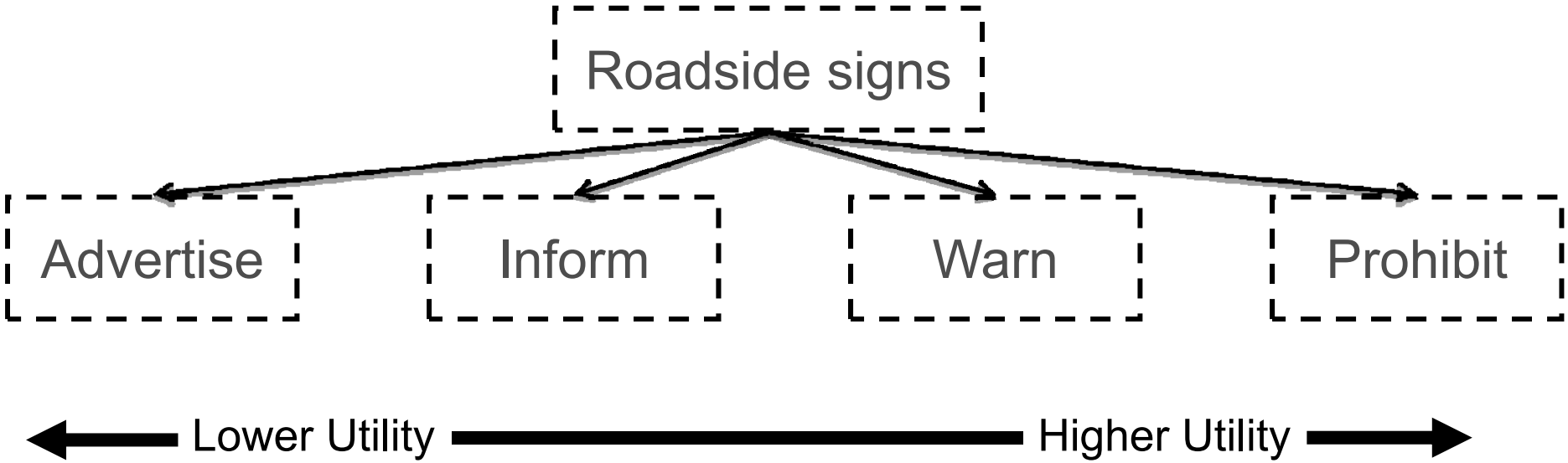
The maly man drank pivo

pi vo

/p/ /i/ /v/ /o/

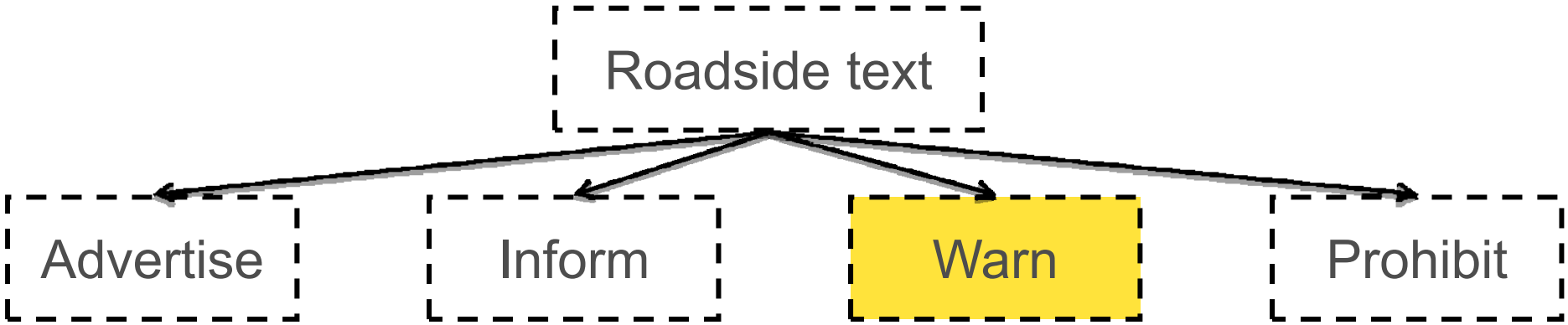
OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example of Utility Application: Roadside Text/Graphics



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example of Utility Application: Roadside Text/Graphics

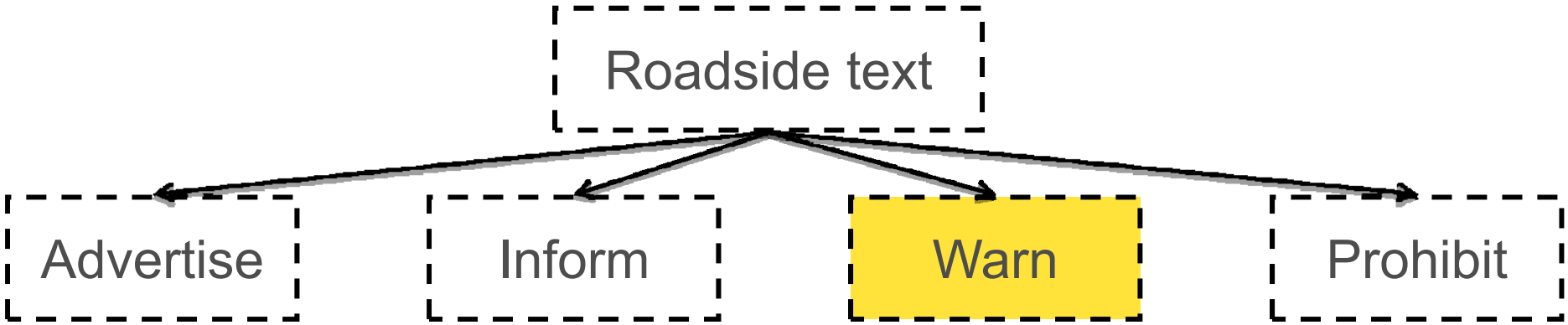




Diamond shape gives the context of the sign
Warning has a high utility

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example of Utility Application: Roadside Text/Graphics

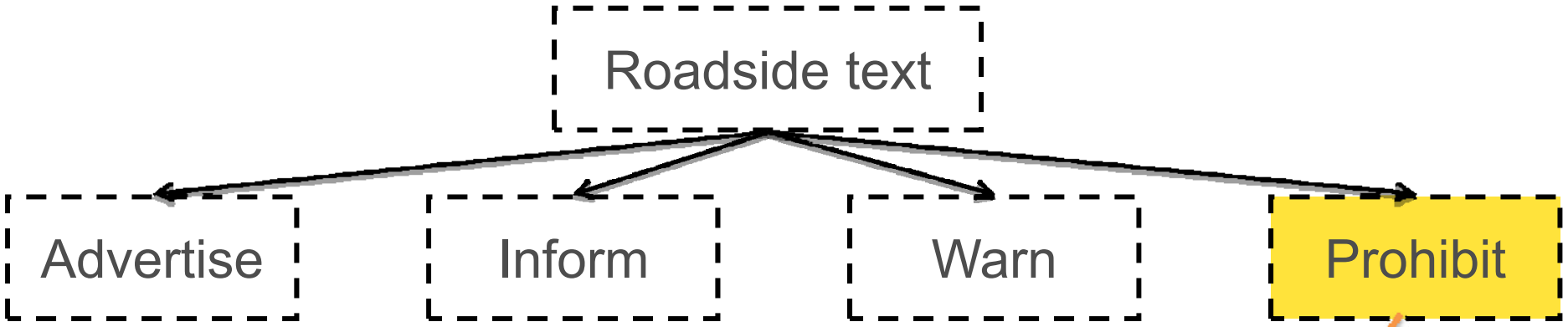




Novel symbol (at some point)
(junction at a bend)

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

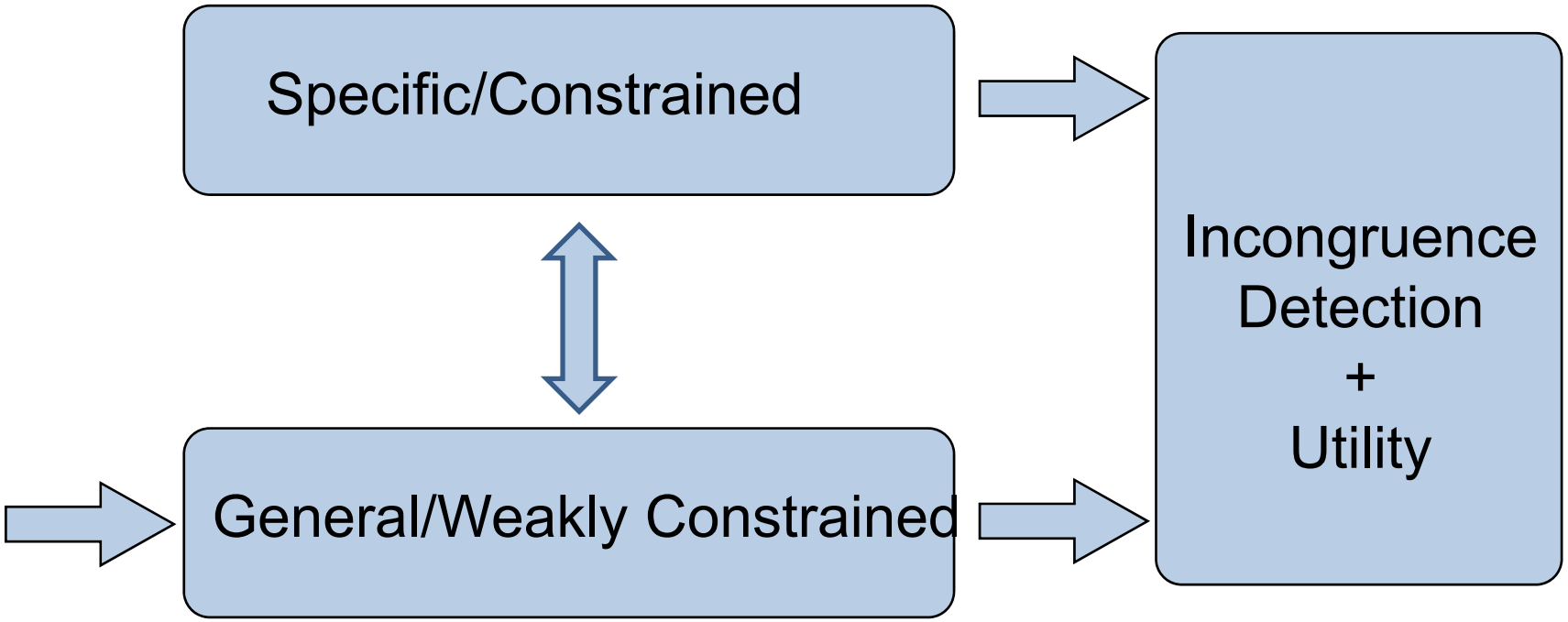
Example of Utility Application: Roadside Text/Graphics





Framework: Detection of Conflict in Probabilities

- Generative approach (a speech example)
- Discriminative approach (a vision example)



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

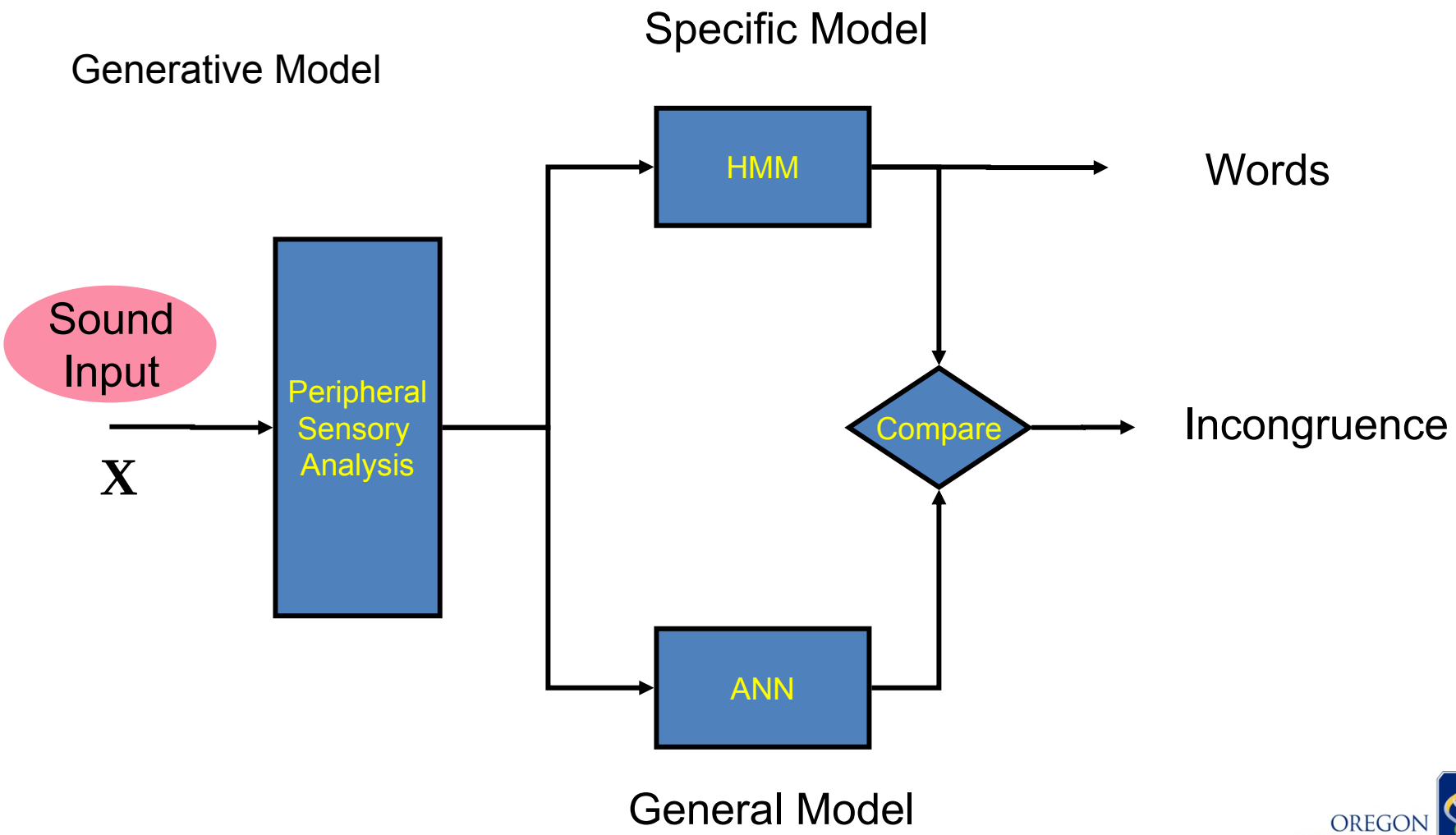
Example: Digit Recognition

Context: “Please say your ten digit account number”

- Expected: Sequence of 10 digits:
- Possible “unexpected” inputs:
  - “O”
  - “Not”
  - “Three hundred twenty”
- Context: “Please say your address”
- Expected: “One six zero zero pennsylvania avenue”
- Possible “unexpected” inputs:
  - “Sixteen hundred pennsylvania avenue”
  - “Pennsylvania avenue sixteen hundred ”

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Example: Detection of out-of-vocabulary (OOV) Words

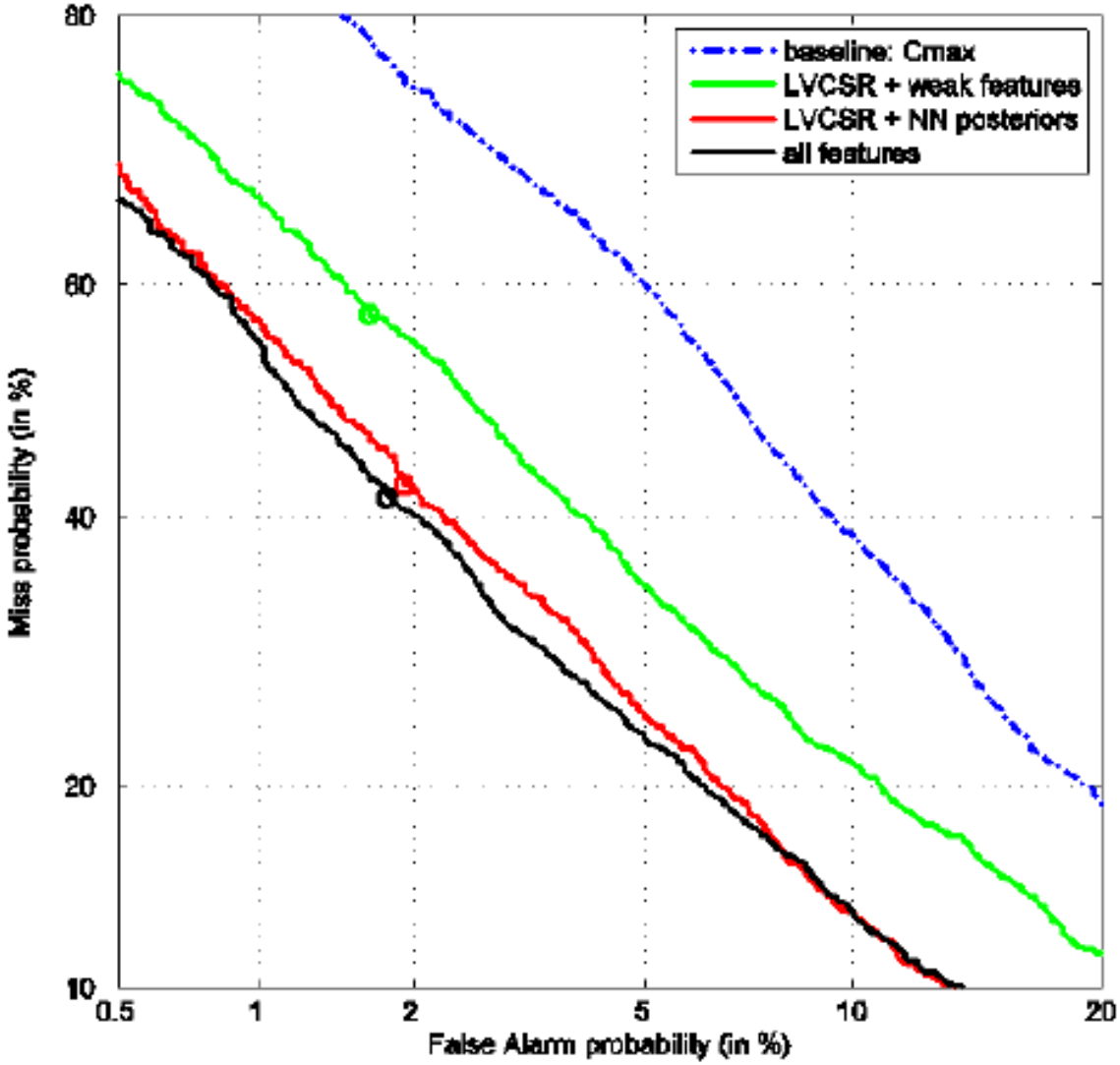


OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Example: Out-of-Vocabulary Word Detection

- Train a spoken digit recognizer on all but one digits
- Test with all digits

Detection Results (without Utility)



Audio-Visual Detection of New Individual

Known Training Identities



Known Testing Identities

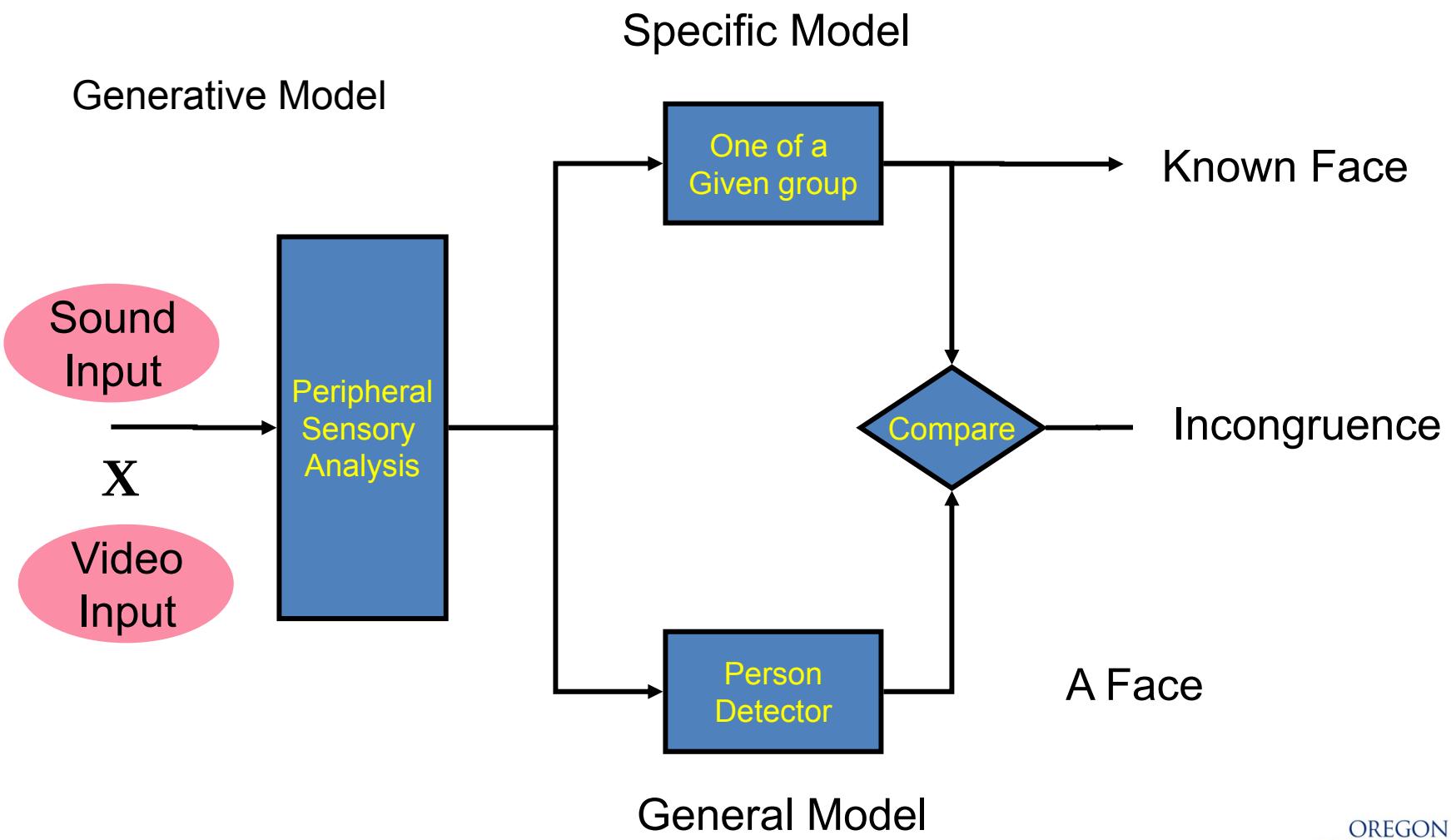


Unknown Identities



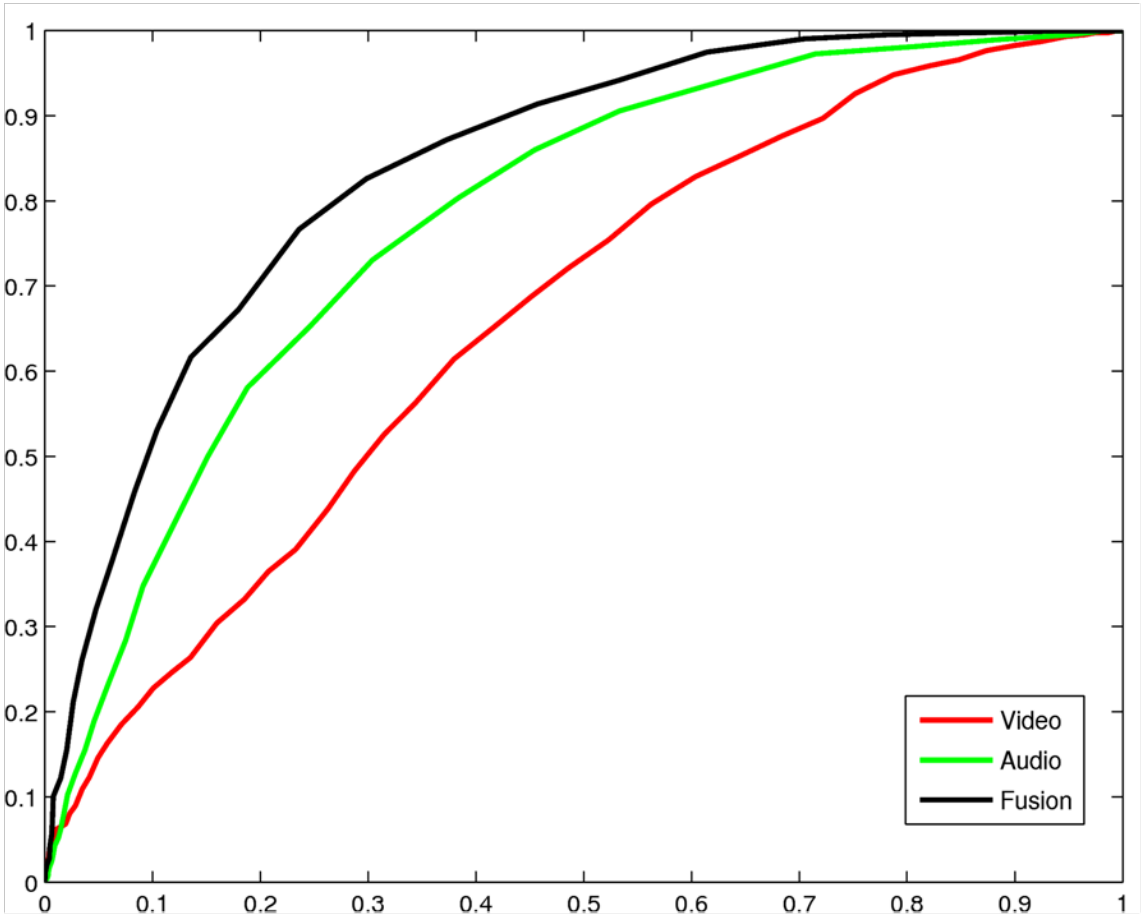


Example: Detection of out-of-vocabulary (OOV) Words



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Results



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Audio-Visual Authentication

- Classify individuals using A/V inputs
- Categories were
  - Face
  - PLP Speech representation



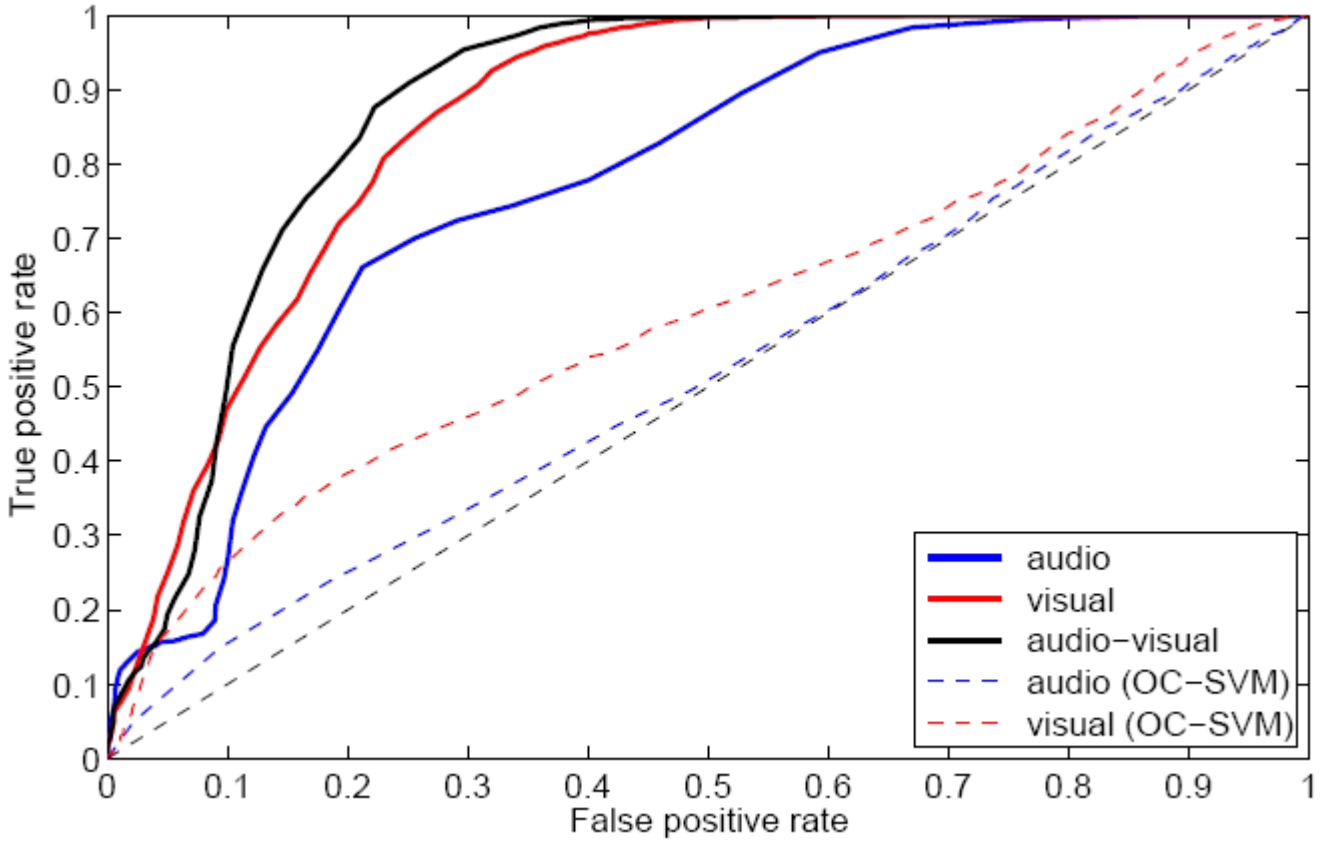
OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Rare – Incongruent Events

	Specific Afgan	General Dog
"Noise" or oddball	Reject	Reject
Incongruent	Reject	Accept
Incorrect Model	Accept	Reject
Not rare or incongruent	Accept	Accept

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

ROC for Imposter Classification

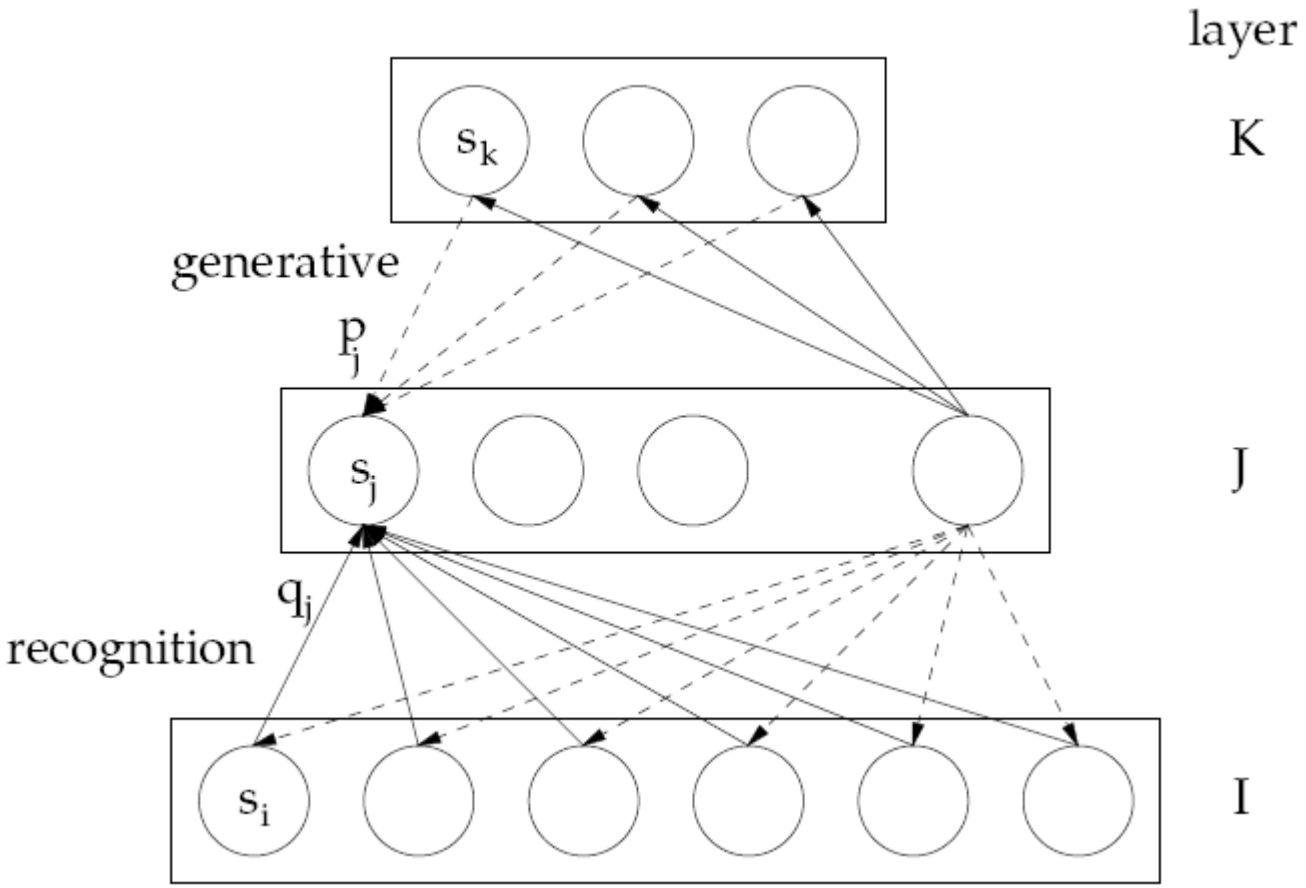


Work in Progress

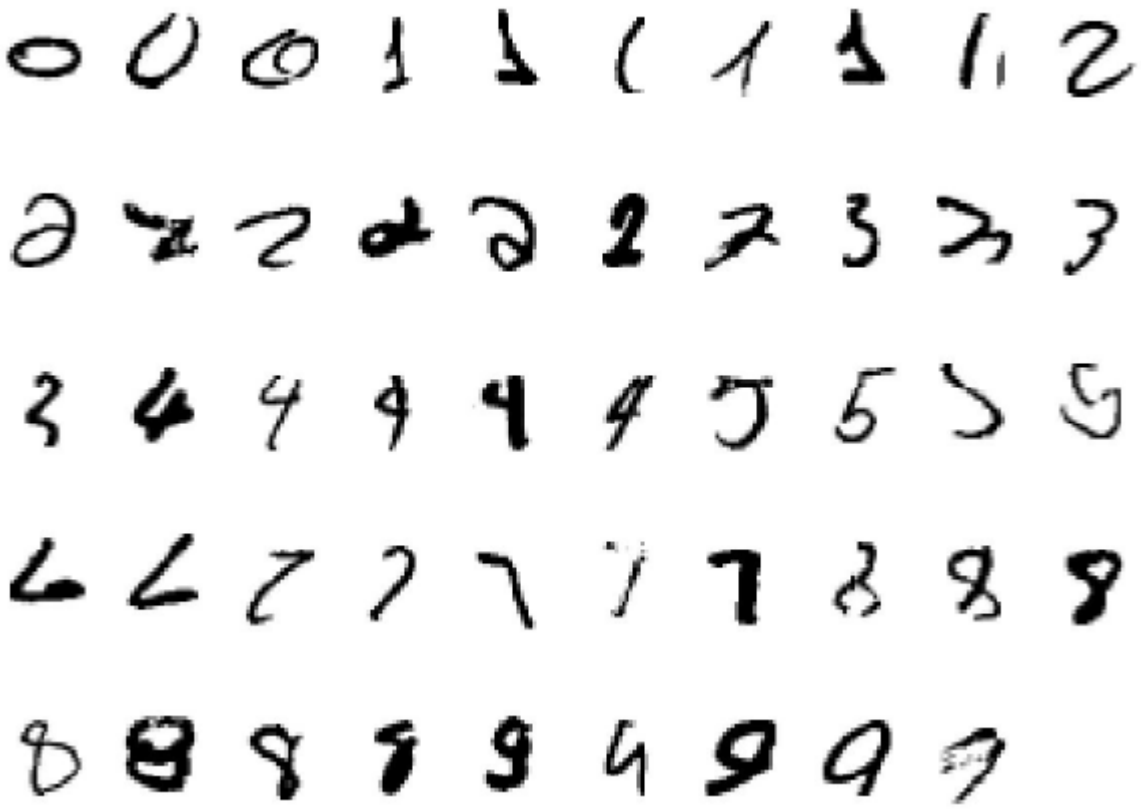
- Using learning by generation (Hinton)
- Generative model of stimulus features
- Leave a digit out

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Learning by Generation (Hinton)



Recognition of Written Digits



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Multimodal Fusion

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

A Number of Mathematical Techniques

- Bayesian
- Mixture of experts
- Bagging, boosting,…
- Theory of evidence (Dempster Schafer)
- Fuzzy set theory

Bayesian Fusion

- Given input X_1, X_2, maximize probability of class C

$$P\{C \mid X_1, X_2\} = P\{X_1, X_2 \mid C\} \frac{P\{C\}}{P\{X_1, X_2\}}$$

- Independence

$$P\{X_1, X_2 \mid C\} = P\{X_1 \mid C\} P\{X_2 \mid C\}$$

- Find class C that maximizes

$$\log\left[\hat{P}\{X_1, X_2 \mid C\}\right] = w_1 f_C(X_1) + w_2 f_C(X_2)$$

- When and how to choose w_i?

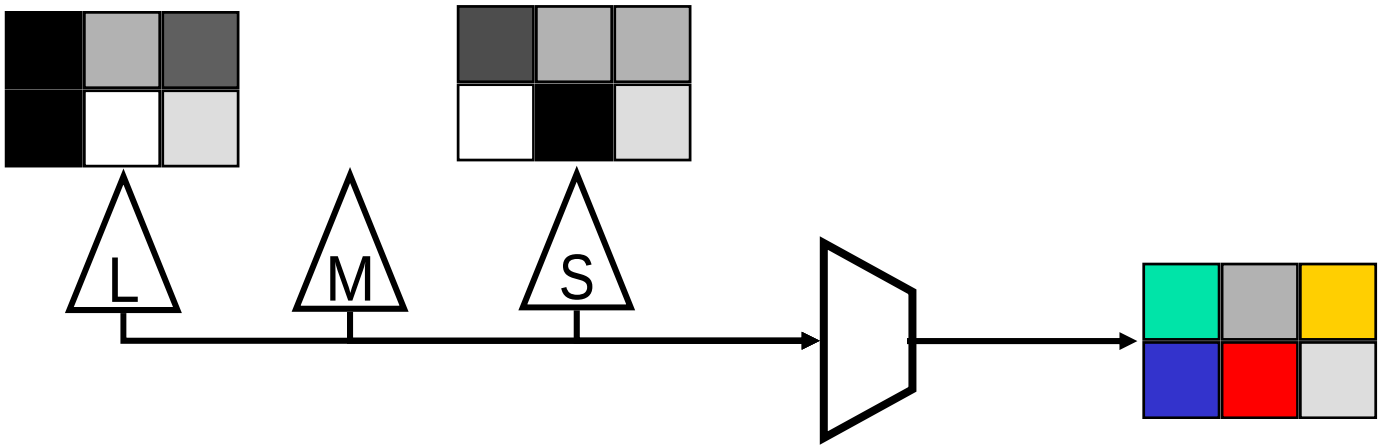
Examples of Fusion in Vision

- Color vision
- Binocular vision and depth perception
- Integration of motion information
- Spatial frequency analysis: multi-scale representation and integration of scales:
  edge location/orientation
- Visual search and object detection

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Color Vision

- Individual sensors (cones) do not see color – How are the sensors combined to form color image?



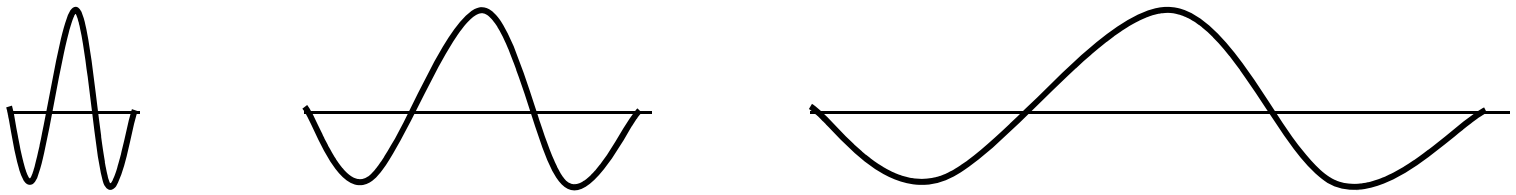
- Color constancy - Integration of color information over space

Robustness: Use context information to perform "white balance" computation

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Spatial Frequency (Scale) Analysis in Vision

- Multi-resolution representation
- Fusion of information from different scale streams – combination of edges

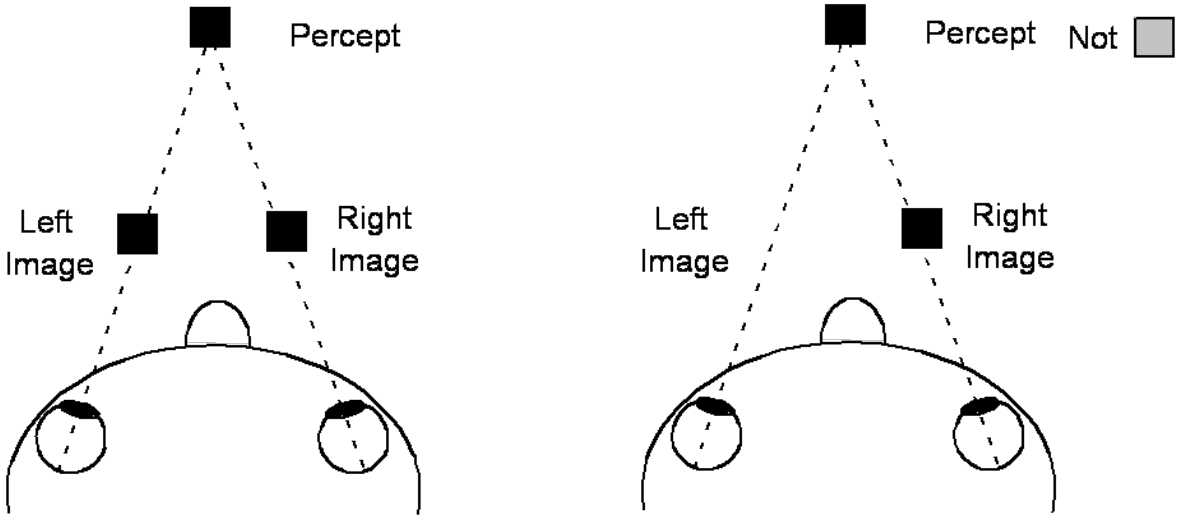


Receptive Fields

- Robustness: Approximate scale invariance
  The impact of a given scale depends on cue reliability

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Fusing Multiple Cues in Binocular Perception (Stereo)



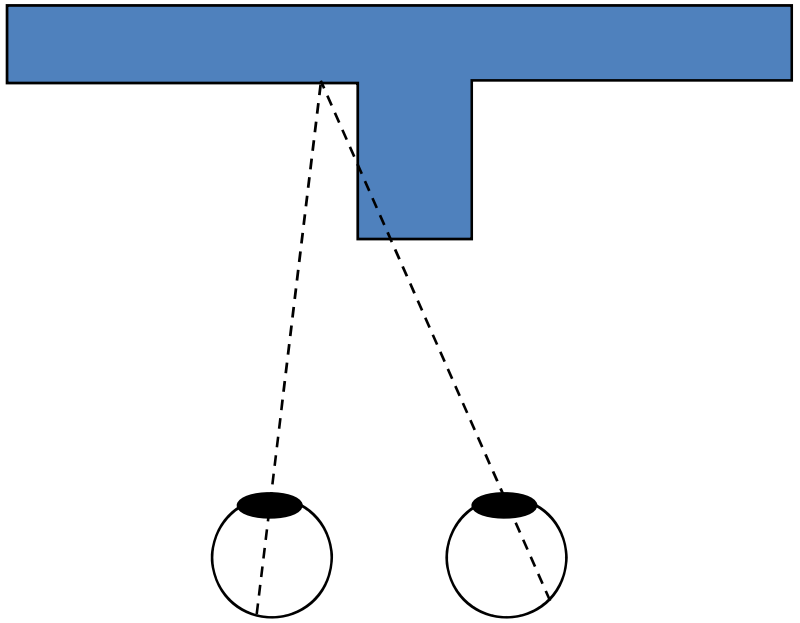
Fusion enables stereo and improves S/N ratio

Fusion is not just linear combination or averaging

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

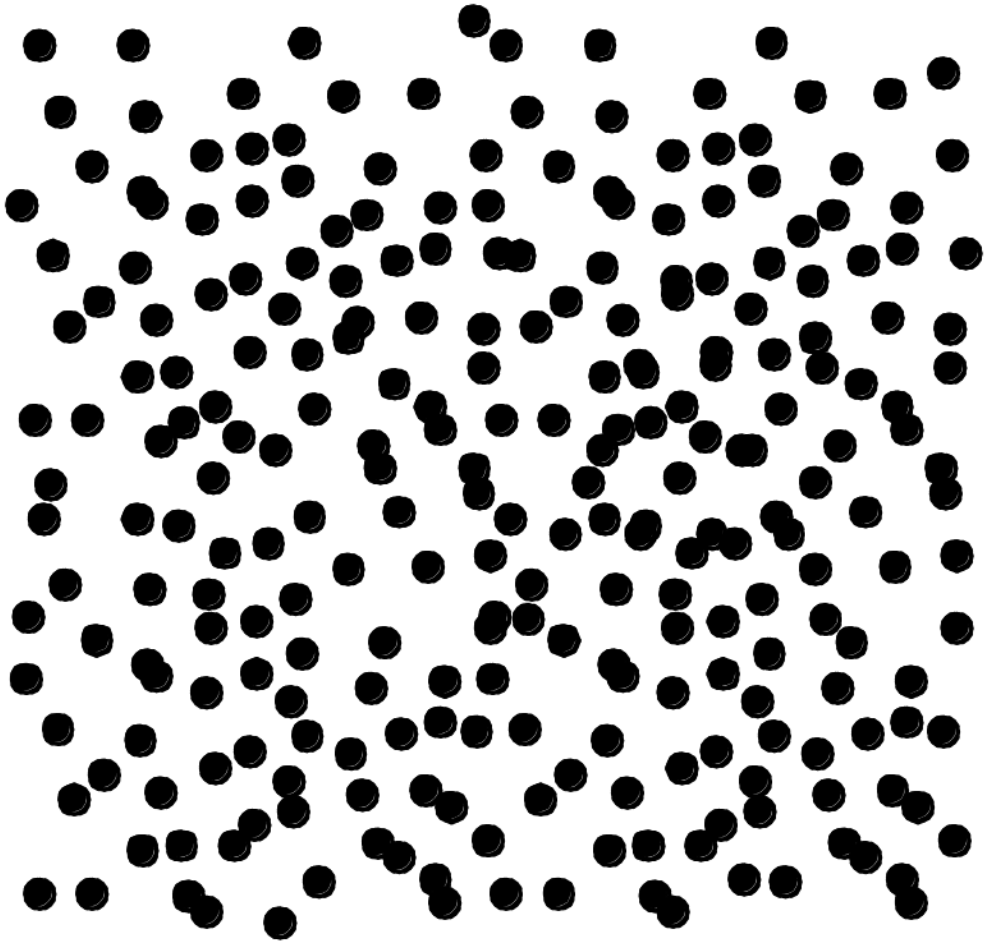
Binocular Rivalry and Da Vinci Stereopsis

- Monocular view due to occlusion
- The “occluded” eye is attenuated



Rapid adaptation – ignore input that is not relevant to a task

Grouping – Motion Integration



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Multiple Cues to Depth Perception

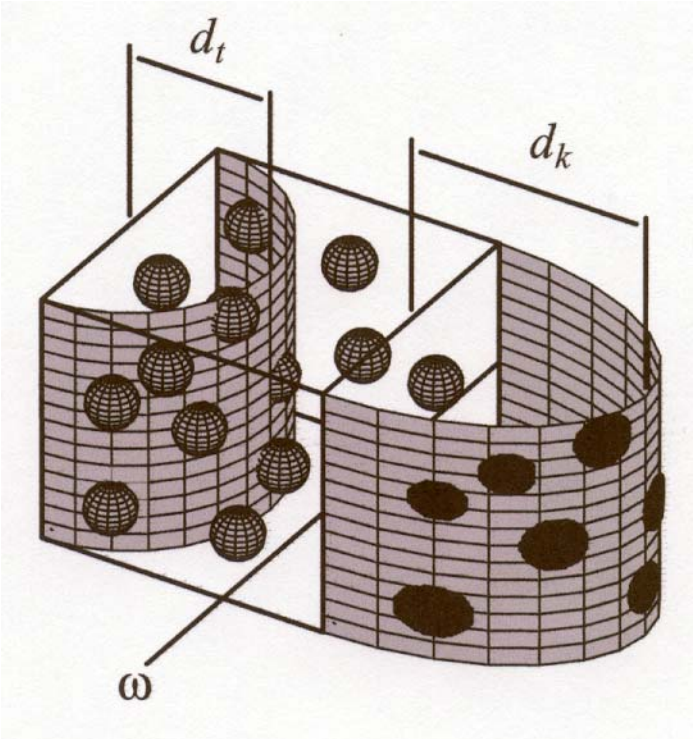
- Perspective
- **Texture**
- Shading
- Stereo by binocular vision
- **Parallax motion** – "Kinetic depth effect"
- Arial perspective (fog, haze)

OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Depth Perception: Fusion of Multiple Cues

- Binocular stereo
- Kinetic depth effect
- Perspective cues

Experiment:

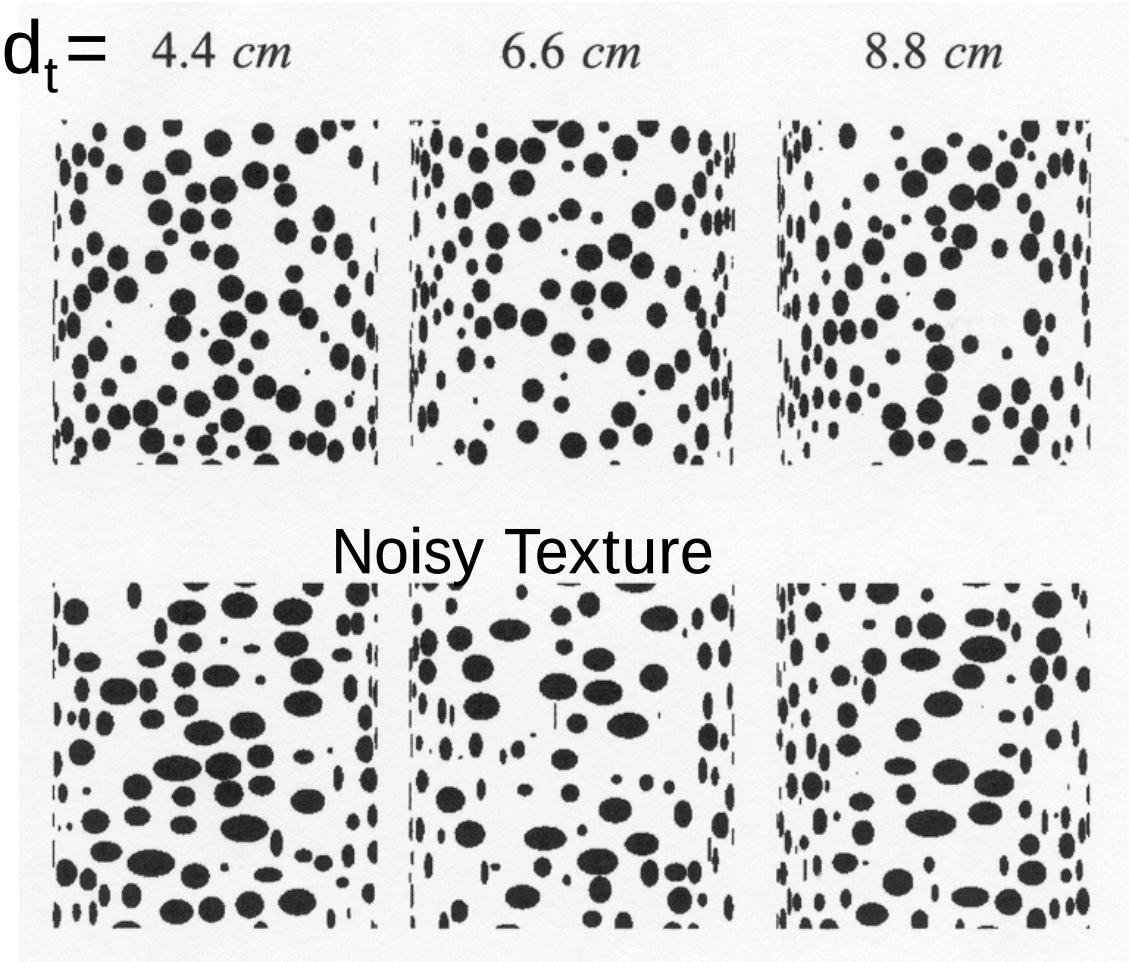


From Landy and Maloney, VR, 1995

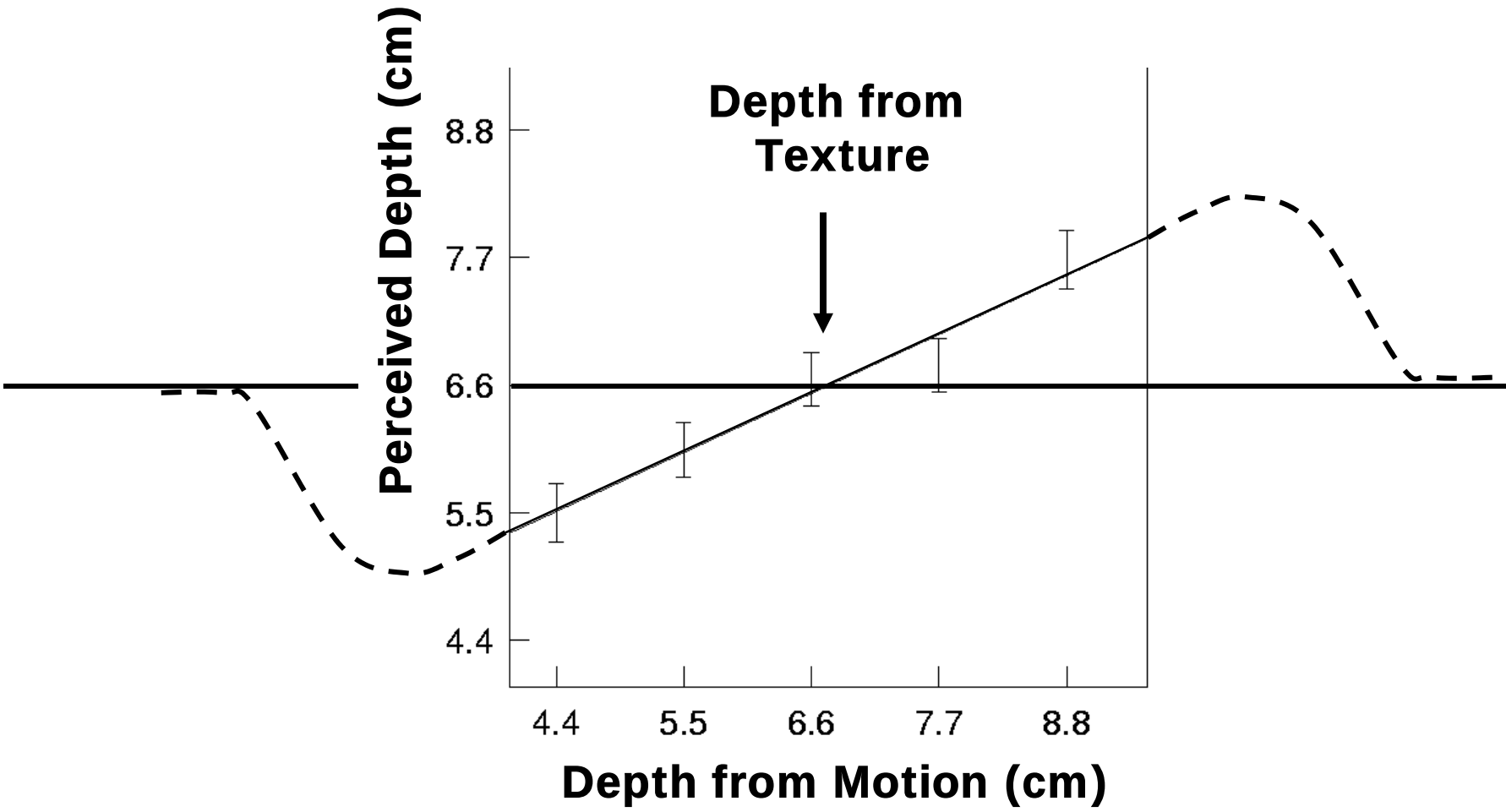
OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Perspective cues

- Texture vs Motion
- Estimate perceived distance
- Result
  $d = w_t d_t + w_k d_k$



Combining Perspective and Motion Cues for Depth Perception

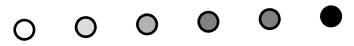


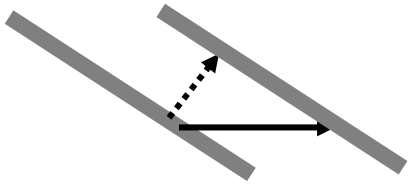
Reliability of information determines the cue weight and ignore input that is not relevant to a task

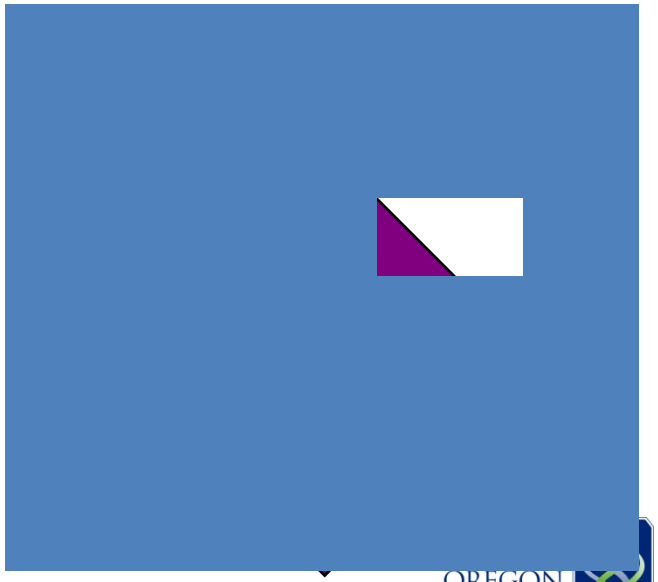
OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Integration of Motion Information

- Object velocity perception
- "Aperture problem"
- Multiple apertures
- Grouping - Affine representation





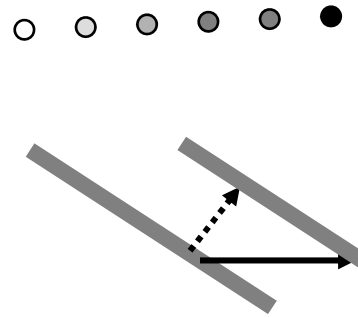


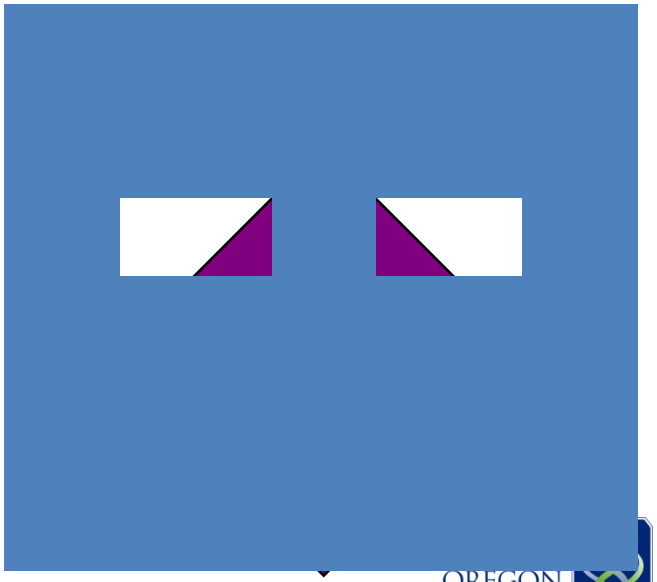
Reliability of local measurements determine their impact on the global estimates

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Integration of Motion Information

- Object velocity perception
- "Aperture problem"
- Multiple apertures
- Grouping - Affine representation





Reliability of local measurements determine their impact on the global estimates

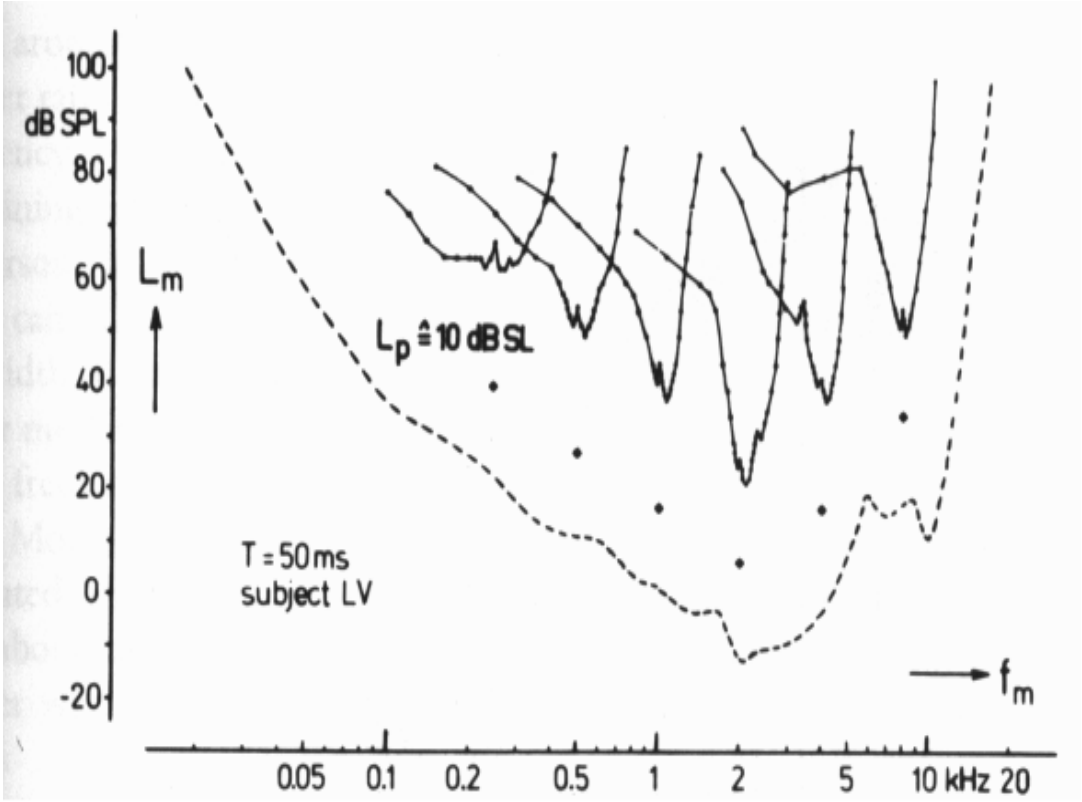
Incorporating Context

- Context ~ Information not directly useful for the classification task, but can modulate performance
- Examples
  - Accent in speech, speaker identity,…
  - Language model X -> {Yes,No}
  - Context in search for objects

Fusion in Audition

- Binaural Hearing
  - Binaural summation – signal detection
  - Binaural release from masking
  - Localization: Source location estimation by cue combination
- Frequency analysis in audition
- Temporal fusion and acoustic stream segregation

Fusing Multiple Streams: Critical Bands Decomposition & Fusion



H. Fletcher's articulation index: Frequency bands appear to provide independent "looks"

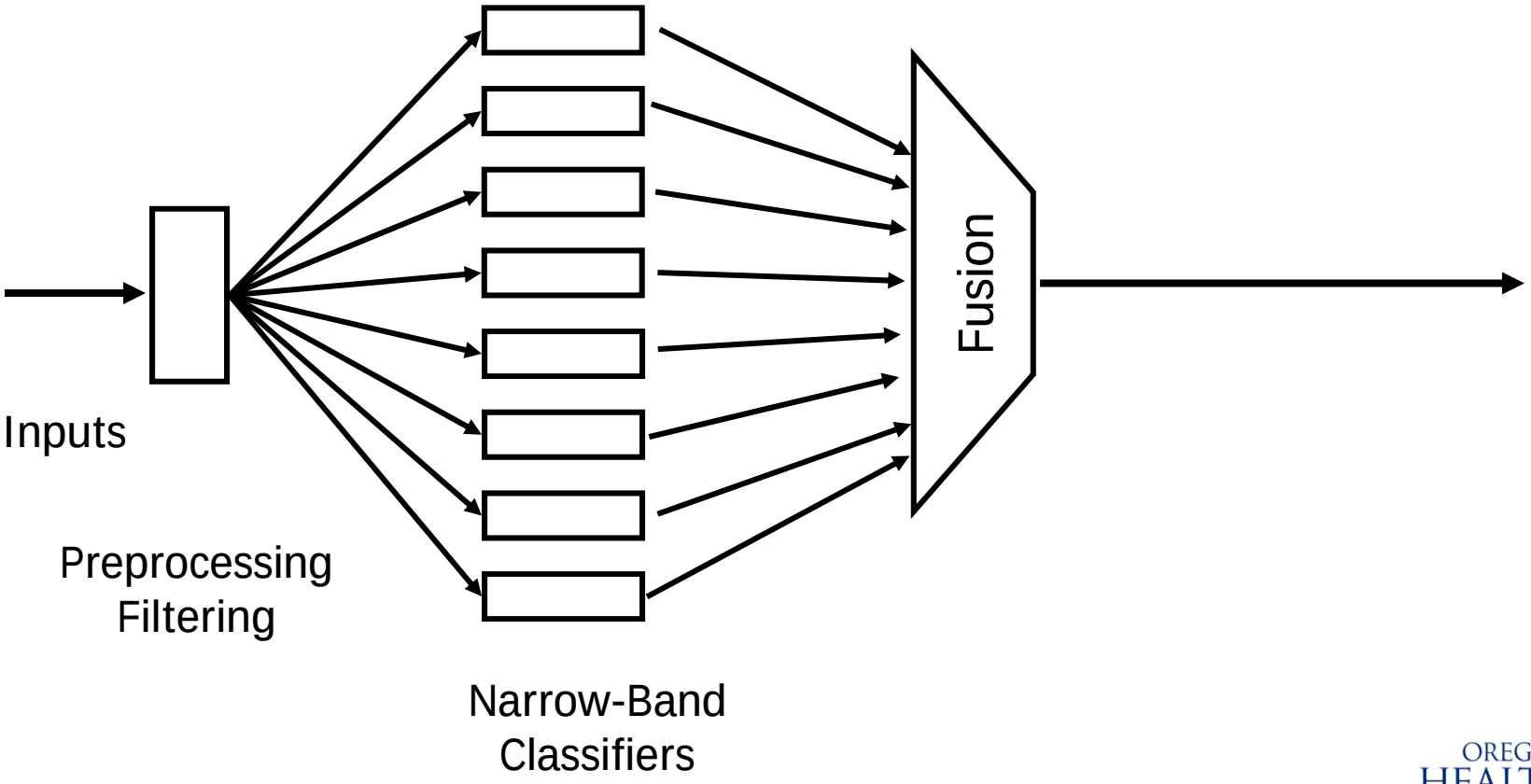
$$P\{Error\} = \prod_B P_B\{Error\}$$

Human speech recognition system ignores unreliable bands?

OREGON HEALTH & SCIENCE UNIVERSITY

Implications of Fletcher's Findings

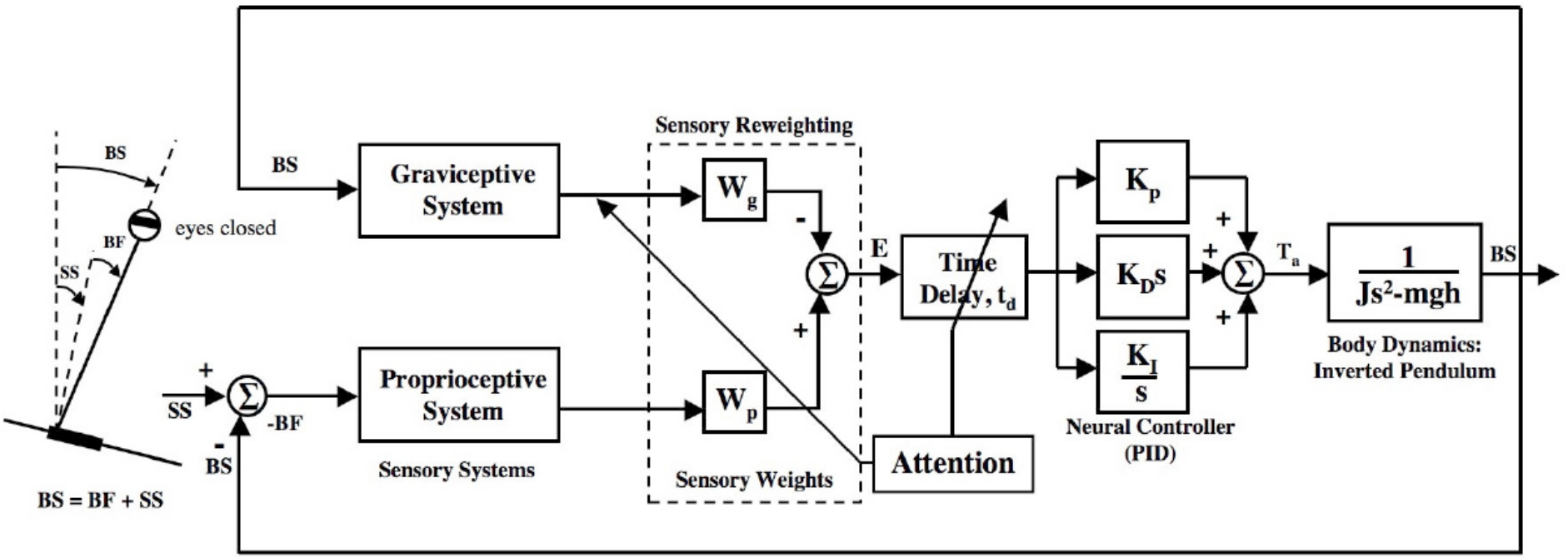
- Independence => No false positives



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Balance and Posture Control

Fusio vestibular, proprioceptive and visual information



OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Fusion in Multimodal Perception

- Visual + auditory speech recognition
  - Complementary information
  - Visual cues override auditory cues
- Gesture + language
  - Dynamic context manipulation
- Synchrony and simultaneity
  - Similar to image registration
  - Tennis game (78 feet)

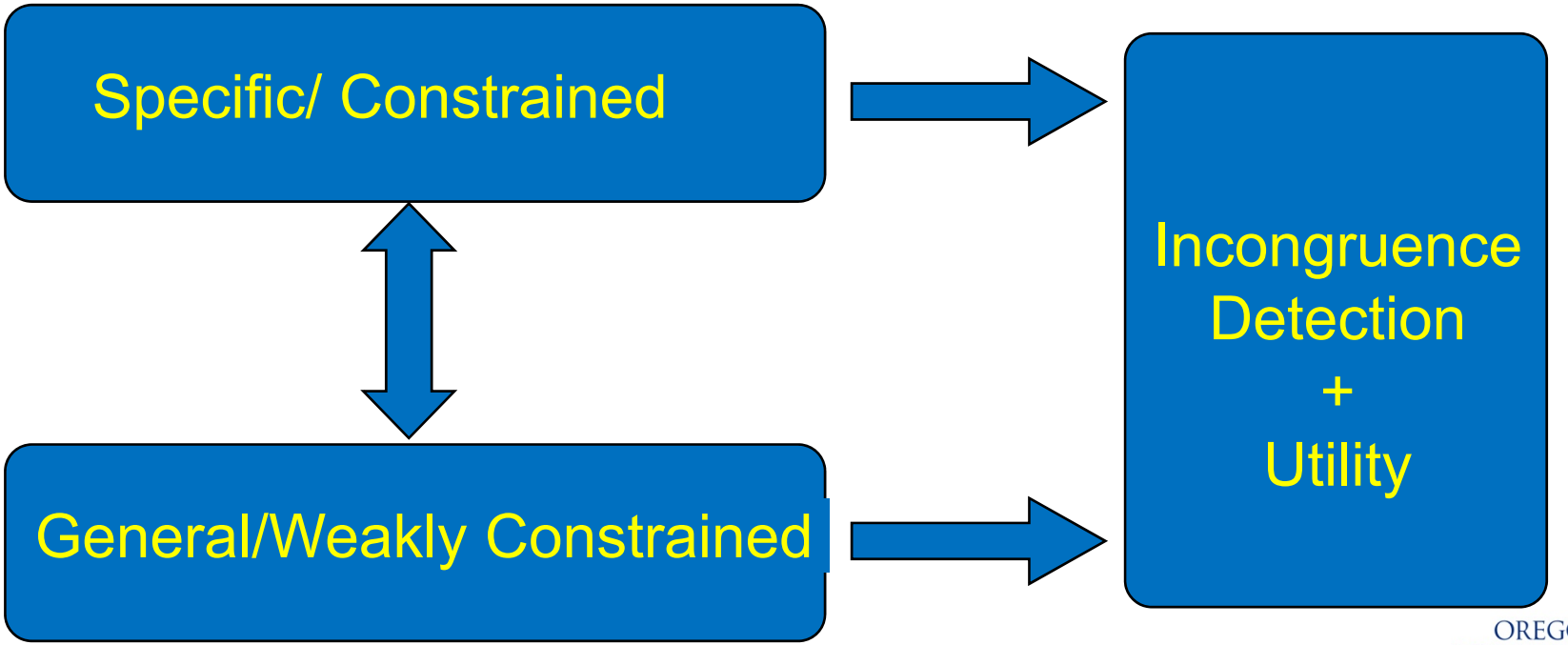
Summary: Benefits of Fusion

Performance enhancements due to

- Improved signal–to-noise ratio
  - Reduction of noise (Registration-conformal mapping)
- Complementary information
  - Visual – auditory recognition (“da” vs. “ba”)
  - Context-aware classification (color vs illumination)
- Higher order features and dimensions
  - Stereo, depth perception (Correspondence)

Summary:

- Detection of Conflict in Probabilities
- Utility of Responses



OREGON HEALTH & SCIENCE UNIVERSITY
OHSU

Application Areas

1. Elder Monitoring
    1. Elders inside/outside activities (falls, near falls, mishaps)
    2. Elders social interactions anomalies
    3. Elders adherence to regiments (medication taking)
2. Surveillance and Security
    1. Analysis of audio/video transmissions
    2. Analysis of interviews
3. Navigation Aids
    1. Navigation in unknown environments
    2. Navigation in support of people with cognitive deficits
4. A/V Appearance Training
5. Deception Detection

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Acknowledgements

- Colleagues and partners in DIRAC

- European Commission Funding of DIRAC
- NIH
- DARPA

OREGON HEALTH & SCIENCE UNIVERSITY OHSU

Thank You

END