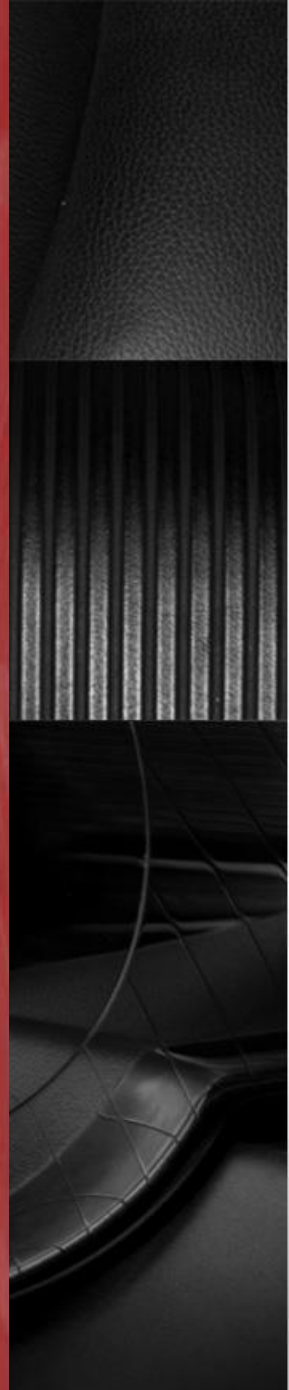# Language Independent Noun And Verb Acquisition From Psychological Videos

By: Diwakar Chauhan

Guide : Amit Mukerjee

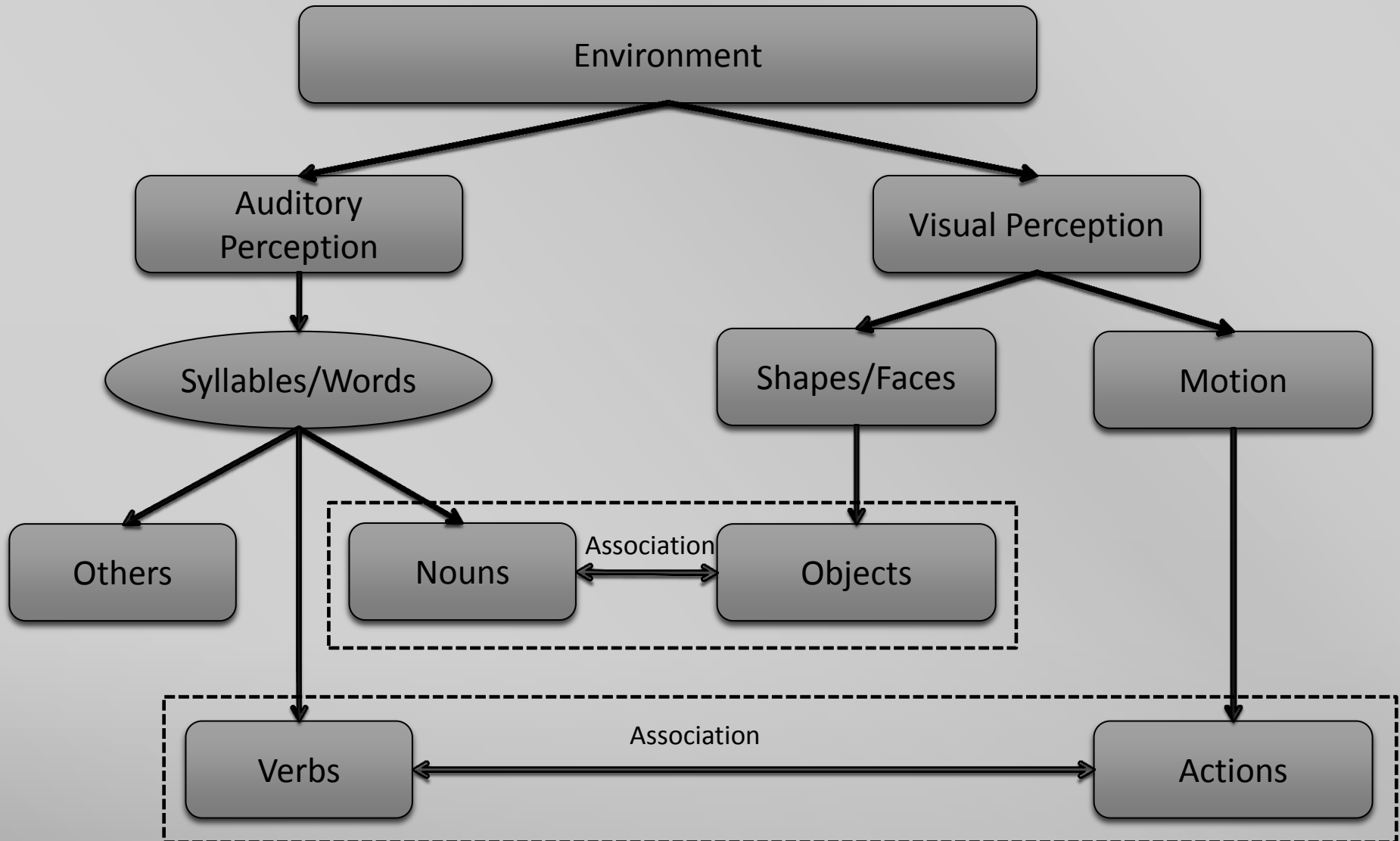email : {diwakarc, amit}@cse.iitk.ac.in
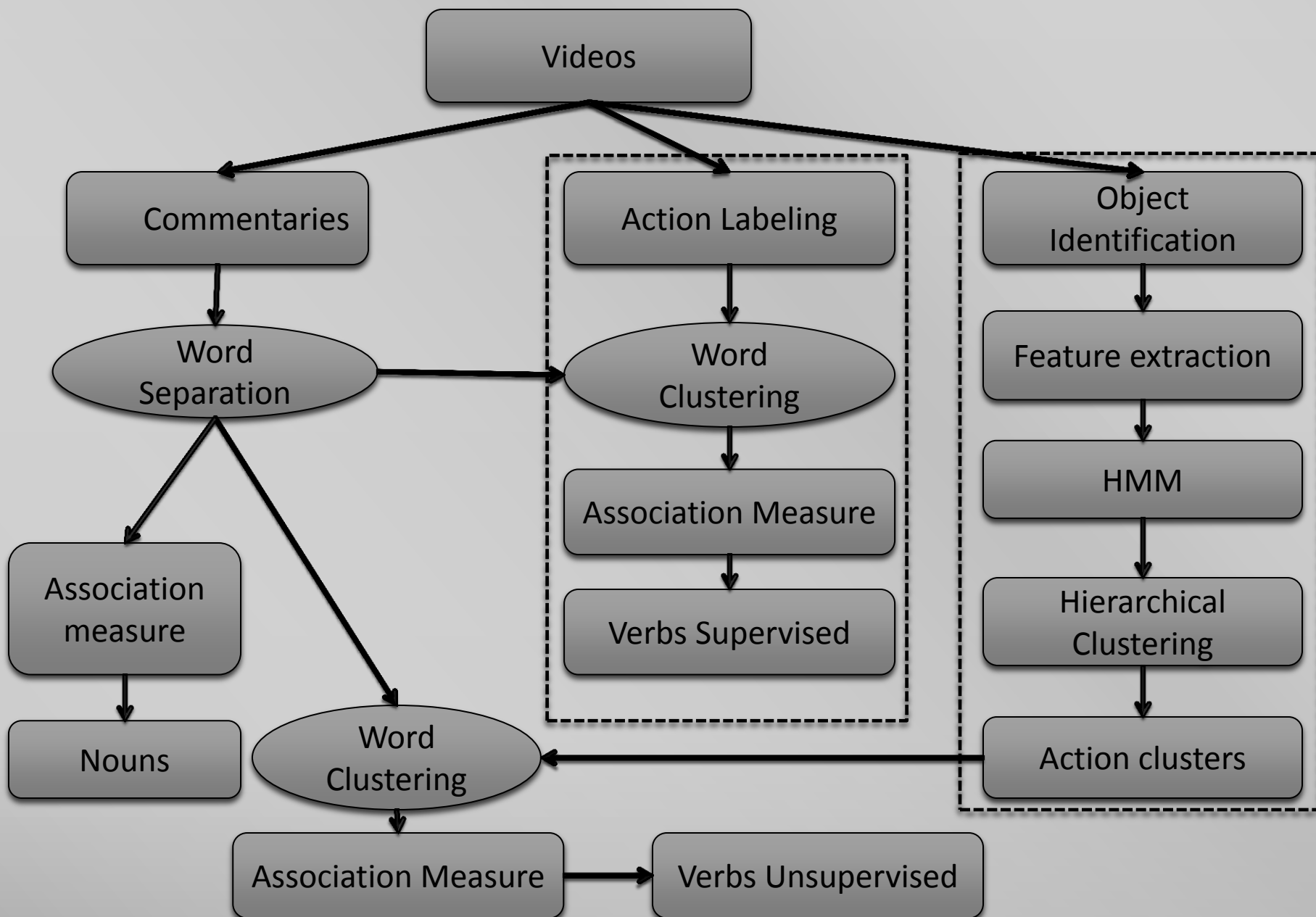
Introduction

- Processing a natural language

  - Requirement of huge resources

  - Applicable to languages with rich corpus e.g. English

  - Can't learn meaning of sentences

- Need of method applicable on multiple languages

- Zero prior knowledge

- Learning process of child
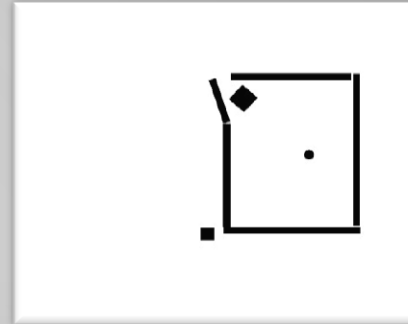
Language learning by Infants

Psychological Videos

- Manually created goal driven videos

- Difficult to generate automatically

- Available Videos

  - Heider Simmel Video

  - Frith Happes Animations
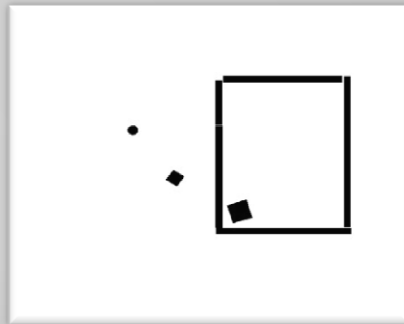
Stills from Heider Simmel Video

Previous Works on Psychological Videos

- G.Satish-Mukerjee, Acquiring Linguistic argument Structure

  - Merge Neural Gas method for action clustering

- Nayak-Mukerjee 2012, Learning Containment Metaphor

  - Learn the language structure and then semantics to get metaphor

Actions in Psychological videos

- Limited actions(chase, go away, come closer) in Heider-Simmel Video

- Need of generalization

- Actions in Frith-Happe Animations

    - Large Variety of actions

    - Chase, push, pull, rotate, play, dance
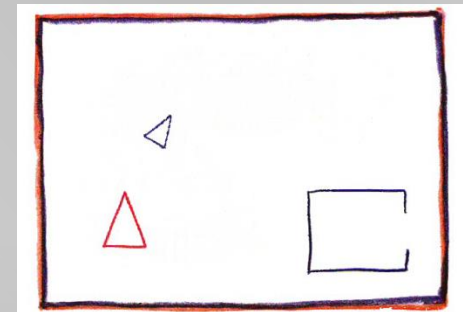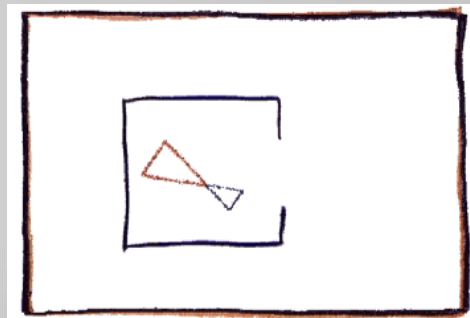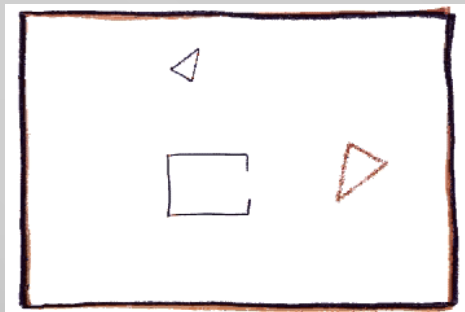
Stills from Frith Happe Animations

Preprocessing of Data and Feature Extraction

- Identification of triangles and Rectangle

  - Line detection : Hough Transform

  - Vertices : Intersection of sides of triangle
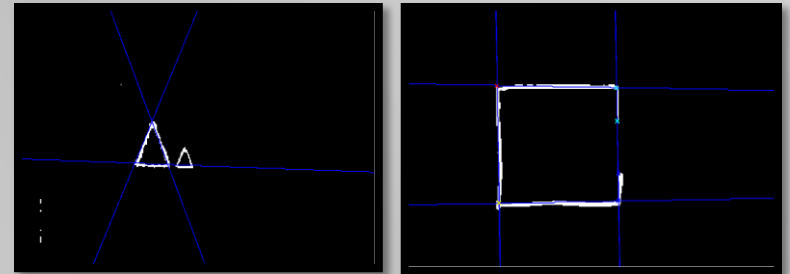
- Feature Extraction

  - Features relevant to motion and interaction of triangles

  - **Centroids of triangles**

    - Captures relative positions of triangles

  - **Orientation of triangles**

    - Capture direction and relative orientations of triangles

Preprocessing of Data and Feature Extraction

- Feature Extraction

  - **Visibility of triangles**

    - Measure of how much one triangle can see other triangle
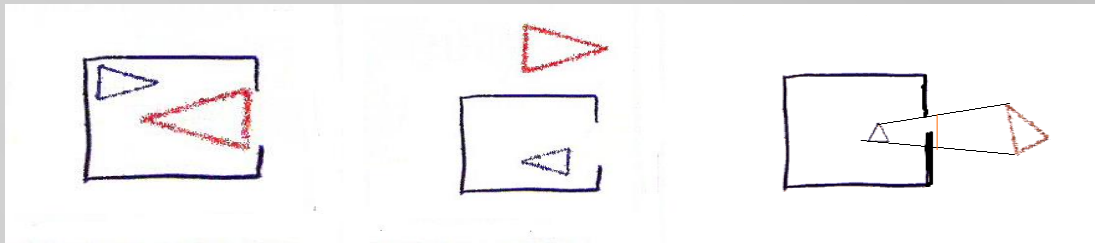
- Feature Vector

  - $\begin{bmatrix} x_1 & y_1 & \theta_1 \\ x_2 & y_2 & \theta_2 \\ & v_{12} & \end{bmatrix}$ or $\quad \begin{matrix} x_1 & y_1 & \theta_1 \\ x_2 & y_2 & \theta_2 \end{matrix} \quad or \quad \begin{matrix} x_1 & y_1 & sin\theta_1 & cos\theta_1 \\ x_2 & y_2 & sin\theta_2 & cos\theta_2 \end{matrix}$

Hidden Markov Models

- Dynamic Bayesian Networks

- Modeled as :-

$$\gamma = \{A, B, \pi\}$$

Where A, B and π are state transition,

observation symbol and initial

probability distribution

- Segment the video into image sequences of some length and train HMM on them
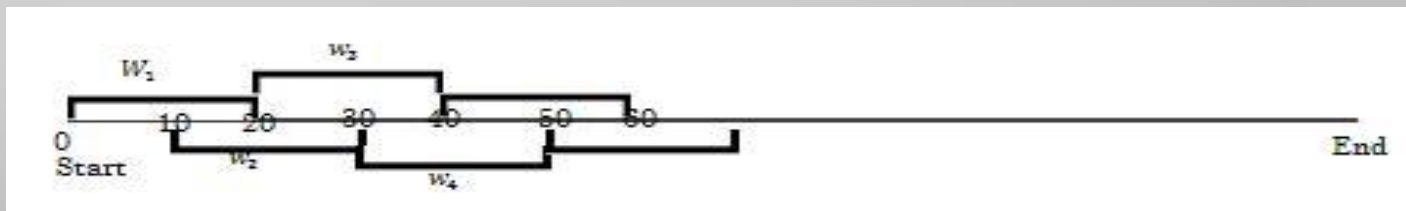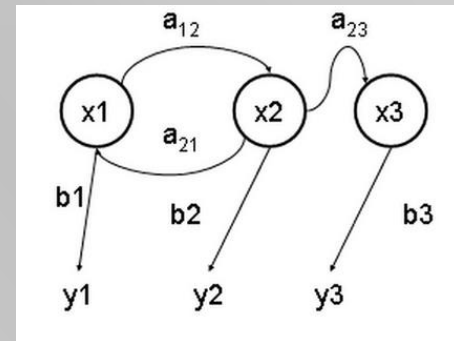
Distance in HMMs and Hierarchical Clustering

- Trained HMMs on all image sequences

- Calculated the Mutual Acceptance measure between HMMs

- Created cluster tree of image segments based on this measure

- Wards method used for cluster distances

- Cut the tree at appropriate height to get 6-8 clusters

- Clusters used for unsupervised verb learning
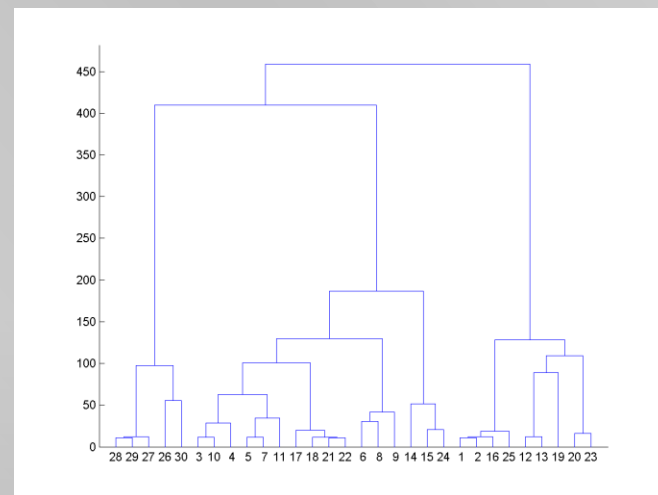
Language

- Lexicons

  - Words or basic units consisting a language (e.g. pitifully)

- Morphology

  - Smallest meaningful elements

  - Word forming from the morpheme (e.g. pity + ful + ly)

- Phonology

  - Syllables (pi + ty + ful + ly)

- Syntax

  - Organization of words to form sentence

- Lexical and phonological properties of a language are considered here

Hindi Corpus

- Obtained Hindi Corpus from CFILT, IIT Bombay

- Corpus has more than 2.8 Million words

- For each unique word in corpus, calculated its frequency and fraction of frequency

- Used it in filtering the equally occurring words in the commentary

- Also Used to remove very common words in results

Commentary Collection and Movement Record

- Collected commentaries in two languages

  - 22 Commentaries in Hindi for Coaxing video

  - 13 Commentaries in English for Coaxing video

- Instructions to subjects before they give commentary

- Converted audio to texts manually

- Merged commentaries of all subjects

- Manually recorded the motion intervals

  of both triangles in video

1	48	एक बक्से के अंदर एक छोटा त्रिकोण एक बड़ा त्रिकोण है
49	76	दोनो आपस में लड़ रहे हैं
77	86	
87	102	और दोनो घूम रहे हैं
103	120	
121	139	छोटा त्रिकोण
140	170	बड़ा त्रिकोण बाहर आ गया बक्से के
171	197	छोटा त्रिकोण अभी भी अंदर है
198	207	
208	238	बड़ा त्रिकोण भी अंदर आ गया अब
239	296	छोटा त्रिकोण बड़े त्रिकोण को बाहर लाने की कोशिश कर रहा है
297	327	
328	374	बड़ा त्रिकोण अब छोटे त्रिकोण को बक्से के बाहर धकेल रहा है
375	400	
401	428	और अब छोटा त्रिकोण बाहर आ गया है
429	460	और बड़ा त्रिकोण अभी भी बक्से के अंदर है

Words and K-Gram Calculation

- Determined all the unique words in the commentaries and their frequency

- Removed common words from the commentary by relative frequency in corpus and in commentary for all words

- Found 1, 2, 3, 4 grams from filtered words

- Determined the frequency and relative frequency of k-grams separately

Noun Learning and Association Measures

- **Definitions**

  - **Concept :**

    - Perception of object or action

    - Concept denoted by c_i and word denoted by w_j

    - Determined by motion records of triangles and simultaneous utterances by subjects

  - Joint Probability

    $$jp\ (c_{i,}\ w_j) = Prob\ (w_j\,|c_i)* prob(c_i) = \frac{time\ when\ concept\ is\ attended\ and\ w_j\ is\ spoken}{total\ Time}$$

  - Mutual Information

    $$mi\ (c_{i,}\ w_j) = jp\ (c_{i,}\ w_j)*\log\left(\frac{jp(c_i,w_j)}{fraction\ w_j\ is\ uttered\ in\ concept\ *fraction\ concept\ is\ attended}\right)$$

  - Relative Frequency

    $$rf(c_{i,}\ w_j) = \frac{Frequency\ of\ w_j\ when\ c_i\ is\ attended}{Total\ frequency\ of\ w_j}$$

Noun Learning

- Concepts used :

    Red triangle is c_1 and blue triangle is c_2

    - When c_1 is moving

    - When c_2 is moving

    - When c_1 is moving but c_2 is not moving

    - When c_1 is not moving but c_2 is moving

    - When both c_1 and c_2 are moving

    - When both c_1 and c_2 are not moving

- Calculated Mutual Information and Joint Probability measures of all k-grams for each of these concepts

User Labeled Action Clustering

- Manually labeled the video with actions

- Total 9 segments of coaxing video labeled

- Learned HMMs on all of these segments

    - Feature vectors were same as that in unsupervised learning

- Calculated Mutual Acceptance measure on all these HMMs

- Merged the clusters with very low HMM distance

- Left with 7 clusters in the end

45	93	घूम रहे, खेल रहे
134	150	गया, बाहर चला गया
151	175	घूम रहा
176	195	अंदर आ रहा
210	230	टक्कर मार रहा, रोक रहा
235	249	ले जा रहा, खींच रहा
310	410	निकाल रहा, ले जा रहा, धकेल रहा, फेक रहा
415	487	बंद कर दिया, खड़ा हो गया, रोक लिया
557	598	घूम रहे, गोल घूमना

Verbs in User Labeled Clusters

- For the intervals labeled, obtained overlapping commentaries

- Selected commentary based on overlap threshold

- Calculated the relative frequency words all words in each cluster/interval

- Filtered the words based on threshold

- This contains very less nouns because nouns are uniform in all clusters/intervals

- Removed the most common words by matching with 1000 most common words in corpus

Verbs in HMM Based Clusters

- Merged all the overlapping clusters from the results of cluster tree cutting

- Removed all very small intervals

  - Very less chance that, they will have actions

- For all the remaining intervals, obtained fairly overlapping commentaries

- Calculated the Relative Frequency measure for all the words in each cluster

- Filtered words based on threshold

- Removed the most common words by comparing them with top 1000 most common words in corpus

Hindi Results : Noun Learning

A. Concept 1(Red Triangle)

■ Monograms

■ Joint Probability

त्रिभुज	**0.245**
बाहर	0.237
लाल	**0.145**
बड़ा	**0.124**

■ Mutual Information

त्रिकोण	**0.510**
बाहर	0.404
बड़ा	**0.335**
त्रिभुज	**0.206**

■ Bigrams

■ Joint Probability

लाल त्रिभुज	**0.100**
नीला त्रिभुज	0.054
छोटे त्रिभुज	0.045
बाहर धकेल	0.041

■ Mutual Information

लाल त्रिभुज	**0.120**
बाहर धकेल	0.117
छोटे त्रिकोण	0.101
छोटे त्रिभुज	0.089

Hindi Results : Noun Learning

C. Concept 1 but Not Concept 2(Strong Red Triangle)

- Monograms
    - Joint Probability

त्रिभुज	**0.034**
बड़ा	**0.029**
त्रिकोण	**0.029**
छोटा	0.021

    - Mutual Information

त्रिकोण	0.158
बड़ा	0.094
छोटा	0.065
टकरा	0.023

- Bigrams
    - Joint Probability

बड़ा त्रिभुज	**0.010**
लाल त्रिभुज	**0.009**
छोटा त्रिभुज	0.008
नीला त्रिभुज	0.007

    - Mutual Information

छोटा त्रिभुज	0.025
बड़ा त्रिभुज	**0.023**
त्रिकोण छोटे	0.021
छोटे त्रिकोण	0.021

English Results : Noun Learning

A. Concept 1(Red Triangle)

- Monograms
  - Joint Probability

triangle	0.165
Smaller	0.127
Outside	0.112
Box	0.104

  - Mutual Information

Triangle	0.211
Trying	0.205
Smaller	0.190
both	0.180

- Bigrams
  - Joint Probability

Smaller triangle	0.100
Bigger triangle	0.054
Small triangle	0.045
Triangle outside	0.041

  - Mutual Information

Bigger triangle	0.094
Smaller triangle	0.087
Inner square	0.072
Small triangle	0.069

Hindi Results : Noun Learning

Blue Triangle : Mutual Information

- Monograms
  - Concept-2

बाहर	0.448
त्रिकोण	**0.430**
छोटे	0.216
त्रिभुज	**0.188**

  - Not Concept-1 Concept-2

नीला	**0.050**
बाहर	0.037
कोशिश	0.032
अंदर	0.028

- Bigrams
  - Concept-2

लाल त्रिभुज	0.151
बाहर निकल	0.100
छोटे त्रिभुज	**0.090**
नीले त्रिभुज	**0.071**

  - Not Concept-1 Concept-2

नीला त्रिभुज	**0.055**
अंदर ना	0.018
लाल त्रिभुज	0.013
वो बाहर	0.013

English Results : Noun Learning

Blue Triangle : Mutual Information

- Monograms

  - Concept-2

Triangle	0.324
Outside	0.291
Smaller	**0.245**
box	0.185

  - Not Concept-1 Concept-2

Outside	0.070
Triangle	**0.066**
Smaller	**0.054**
Success	0.051

- Bigrams

  - Concept-2

Smaller triangle	0.186
Bigger triangle	0.107
Red triangle	0.050
Blue object	**0.041**

  - Not Concept-1 Concept-2

Gets success	0.051
Red triangle	0.051
Ring basically	0.040
Cannot enter	0.040

Hindi Verbs from User labeled clusters

- Ground Truth

- Results

Interval	Verbs
45-93, 151-175,557-598	घूम रहे, खेल रहे, घूम रहा, गोल घूमना
134 - 150	गया, बाहर चला गया
176 - 195	अंदर आ रहा
210 - 230	टक्कर मार रहा, रोक रहा
235 - 249	ले जा रहा, खींच रहा
310 - 410	निकाल रहा, ले जा रहा, धकेल रहा, फेक रहा
415 - 487	बंद कर दिया, खड़ा हो गया, रोक लिया

Verbs
खुश, **पकड़**, गोल, **घूम**, कोण, दोनो, दोस्त
घूमे
वापस
निकाल, **धकेल**, **फेक**, तरीक़ो, **धक्का**, सिरे, जुड़, **धक्के**, **मारकर**, अधूरे, **धकेलने**, चतुर्भुज
रास्ता, **खड़ा**, युध, दरवाजा, ना, भगा, विरूध, पाए, पड़ा, वो, **रोक**

English Verbs from User labeled clusters

- Ground Truth
- Results

Interval	Verbs	Verbs
45-93, 151-175,557-598	Circle, rotate, play	Fighting, **playing**, interacting, **circling**, touching, connected, enjoying, enclosed
134 - 150	Go away	Chasing, dancing
176 - 195	Come	Still
210 - 230	Hit, stop	Funny, ways, corner
235 - 249	Pull	
310 - 410	Throw, push	**Pushing, Push**, Win, wait, moving, ahead
415 - 487	Block, stand, roam	**Blocks**, completely, explore, **roam**, reason

Hindi Results : Unsupervised Verb Learning

- Ground Truth

Interval	Expected
190-255	आना, टकराना
150-195	घूमना, आना
520-595	घूमना
370-405	धक्का
400-455, 460-515	बंद करना

- Results

Verbs Discovered
भिड़ा, हिचक, निकालना, चिढ़ा, शरमा, खींचने, **टक्करमारने**
डरता, खेलना, **घूमे, घूमते**
वापस
युध,

Conclusion and Further work

- The nouns are successfully discovered

- Verbs obtained in user labeled clusters are vey much same as the ground truth

- In the unsupervised learning, the verbs are not mapped very well to action

  - Clustering by HMM

  - Changing time lag in commentary by each subject

- Use of gaze data for building attention model

- Syntax of the language can also be learned after verbs are discovered

Thank You!

Questions?