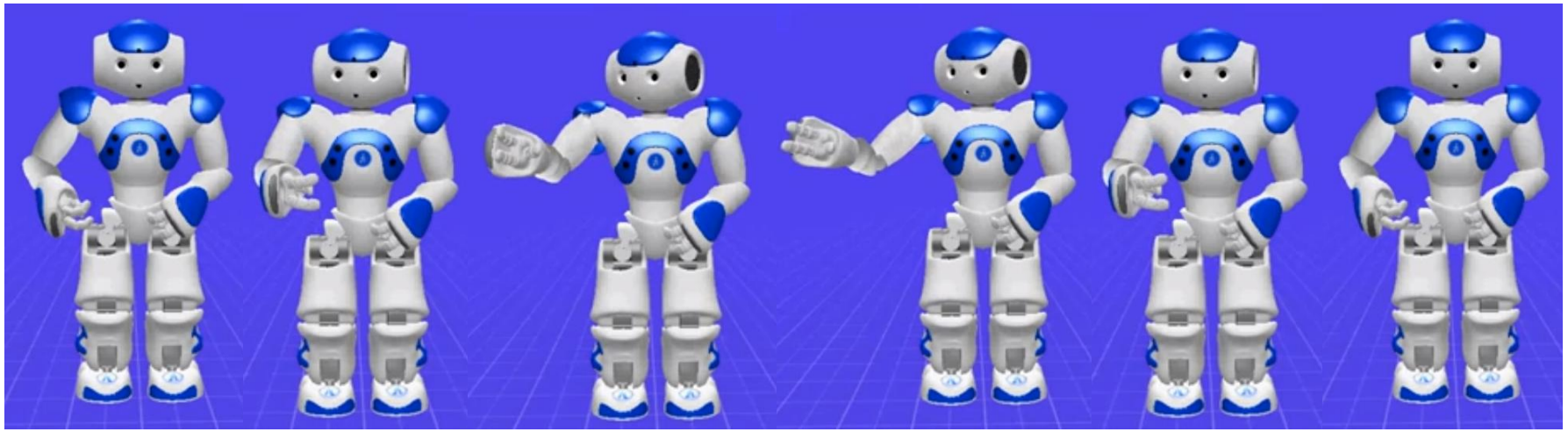
A Grounded Framework for Gestures and its Applications

Debidatta Dwibedi
Mentor: Prof. Amitabha Mukerjee

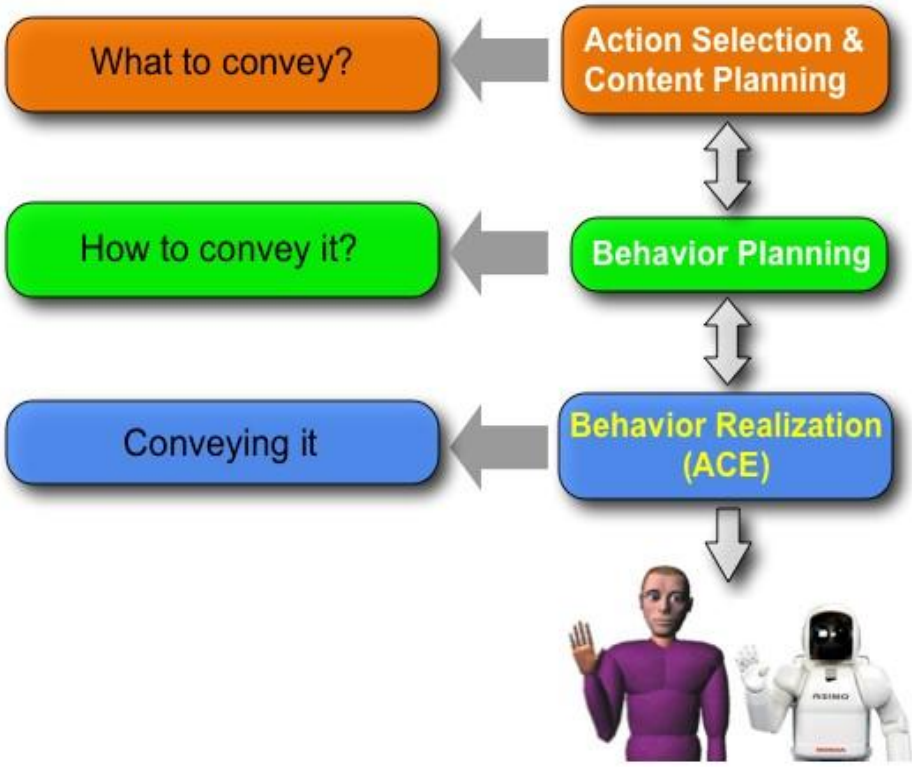
- **Problem Description**
  - Implement a framework that is able to associate gestures with language
  - Generate gestures helpful in route descriptions and assembly instructions

Keywords: Gesture Recognition, Natural Language Processing, Embodied Cognition, Symbol Grounding, Lifelong Learning



Motivation and Past Work

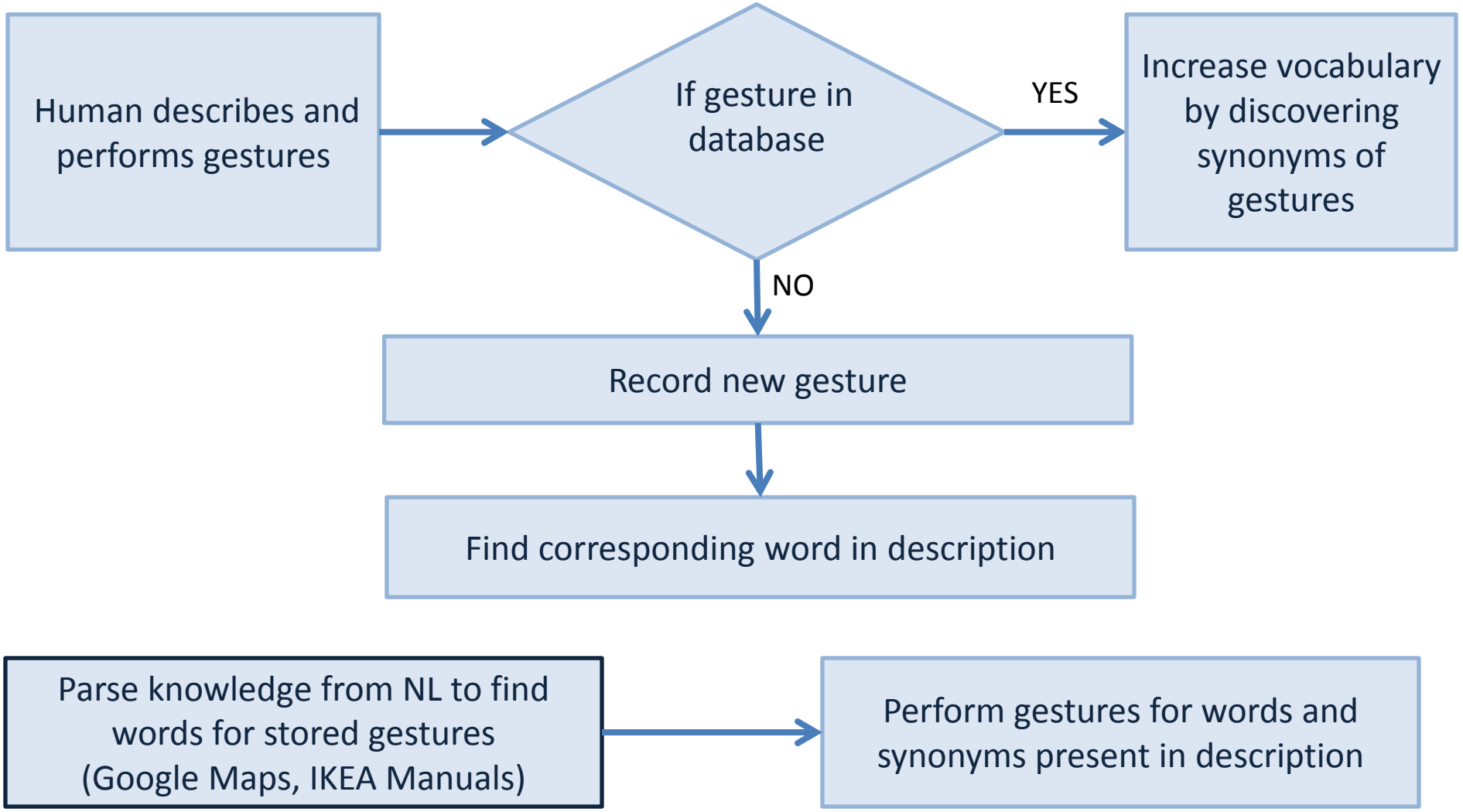
- Knowledge representation for generating locating gestures in route directions ***[Striegnitz 2005]***
- Integrated model of speech and gesture in robots ***[Kopp 2008]***
- Studied models of utterance, gesture and timing to better facilitate the human-robot interaction ***[Okuno 2009]***
- A joint model of language and perception for grounded attribute learning ***[Matuszek 2012]***



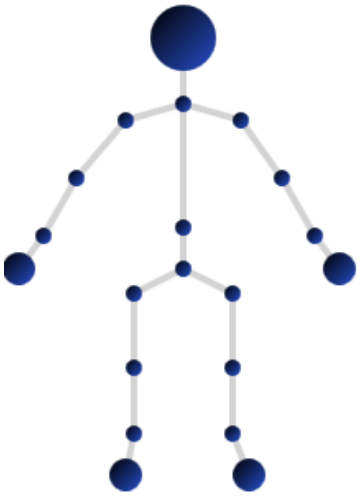
Kopp et al 2008

Towards a Lifelong Gesture Learning System

Framework for Lifelong Learning of Gestures



Recording and Recognizing Gestures



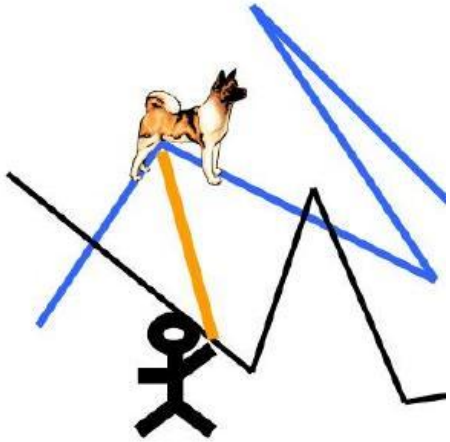
http://soulsolutions.com.au

Get coordinates of :

- (LeftHand - HipCenter) [x_1,y_1,z_1]
- (RightHand - HipCenter) [x_2,y_2,z_2]

This 6 X N dimensional vector is the representation of the gesture,

where N is the no. of frames taken to do the gesture i.e. time



http://www.cs.tulane.edu/~carola/research/mandog2.jpg

Used Frechet distance between query gesture and gestures in database to find nearest match.

If distance exceeds a threshold, then gesture is assumed to be not present in database.

Associating Words with Gestures

Collected 2 datasets and did frequency analysis:

- **Route descriptions in IITK**
  - right, left , turn, straight, hall, road, walk, building ,('take', 'right'), ('take', 'left'), ('go', 'straight'), ('turn', 'left'),('turn', 'right'), ('right', 'turn')
- **Assembly of a TV stand**
  - shelf, frame, glass, place, top, bottom, bolts , ('allen', 'wrench'), ('shelf', 'frame'), ('bottom', 'shelf') , ('glass', 'shelf'), ('top', 'shelf')

Gesture for Word	Right	Left	Straight	Building
Right	**69000**	0.232	0.189	0.045
Left	0.27	**61000**	0.196	0.0517
Straight	0.458	0.399	**35000**	0
Building	0.272	0.272	0.166	**14000**
Turn	1.624	1.333	0.199	0.049
Walk	1.222	0.666	0.249	0.111
Go	0.499	0.333	0.919	0.021

Transferring Gestures and Words to ECA

Nao robot is chosen as ECA(Embodied Cognition Agent)

- Choreographe to simulate on an animated version of Nao.
- Gestures easily transferrable to the actual robot.

Recorded all 20 joints from Kinect

- But joints in Cartesian coordinates and Nao takes angles as input.
- Inverse kinematic problem.

Converted into corresponding joint angles

Can utter instructions from Google Maps or IKEA manuals

- In-built text-to-speech engine

Timing and spacing of instructions and gesture performance has a major impact of comprehension of gestures

Conclusions and Future Work

- **An important part of the system is the module that parses knowledge from systems like Google Maps to find what are the gestures that can be performed.**
- **Integrate all the modules into one lifelong gesture learning system.**
- **Use of gestures in a collaborative setting.**
- **Use speech input to find association between gestures and words.**
- **Carry out usability studies**
  - Impact of orientation of robot, timing of utterances and gestures etc. on comprehension
  - Provide gestures and descriptions to humans and see effectiveness of such a system

References

[Striegnitz 2005] Knowledge representation for generating locating gestures in route directions
[Kopp 2008] Multimodal communication from multimodal thinking—towards an integrated model of speech and gesture production.
[Okuno 2009] Providing route directions: design of robot's utterance, gesture, and timing.
[Matuszek 2012] A Joint Model of Language and Perception for Grounded Attribute Learning.

Acknowledgements

I would like to thank **Nehchal Jindal** and **Shashank Sonkar** for their valuable inputs in this project. I would also like to acknowledge **Prof. Mukerjee** for all the guidance and support he has provided throughout this semester.