Assistive Robots: Teaching Asimo ASL

Xiao Ge
Shruti Gupta
Taylor Newton

Two main steps

- Integrate Voice Recognition Software
  - We did this by using a program called Sphinx which wrote the input into a text file which we read from our program.
- Generate trajectories
  - We generated the trajectories in joint space because of the precise angle of the letters in American Sign Language.

Assistive Robots

University of South California

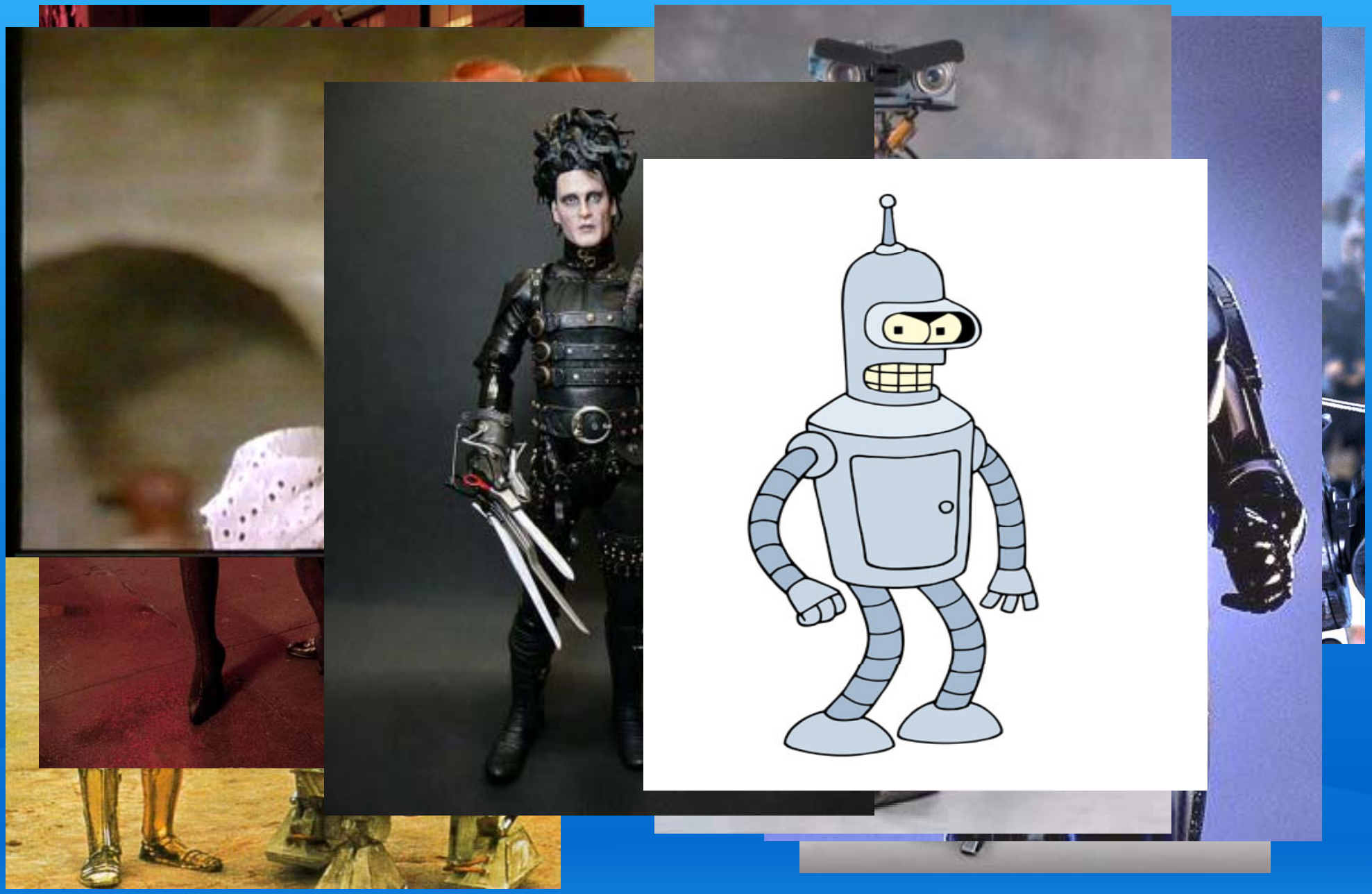
Assistive Robots

MIT: MIT postdoctoral associate Aaron Edsinger gets some help from Domo, an assistive robot he has been developing for the last three years. (circa 2007)

Previous Assistive Robotics Projects

http://humanslab.ece.gatech.edu/Projects/PlayingPalsy.html

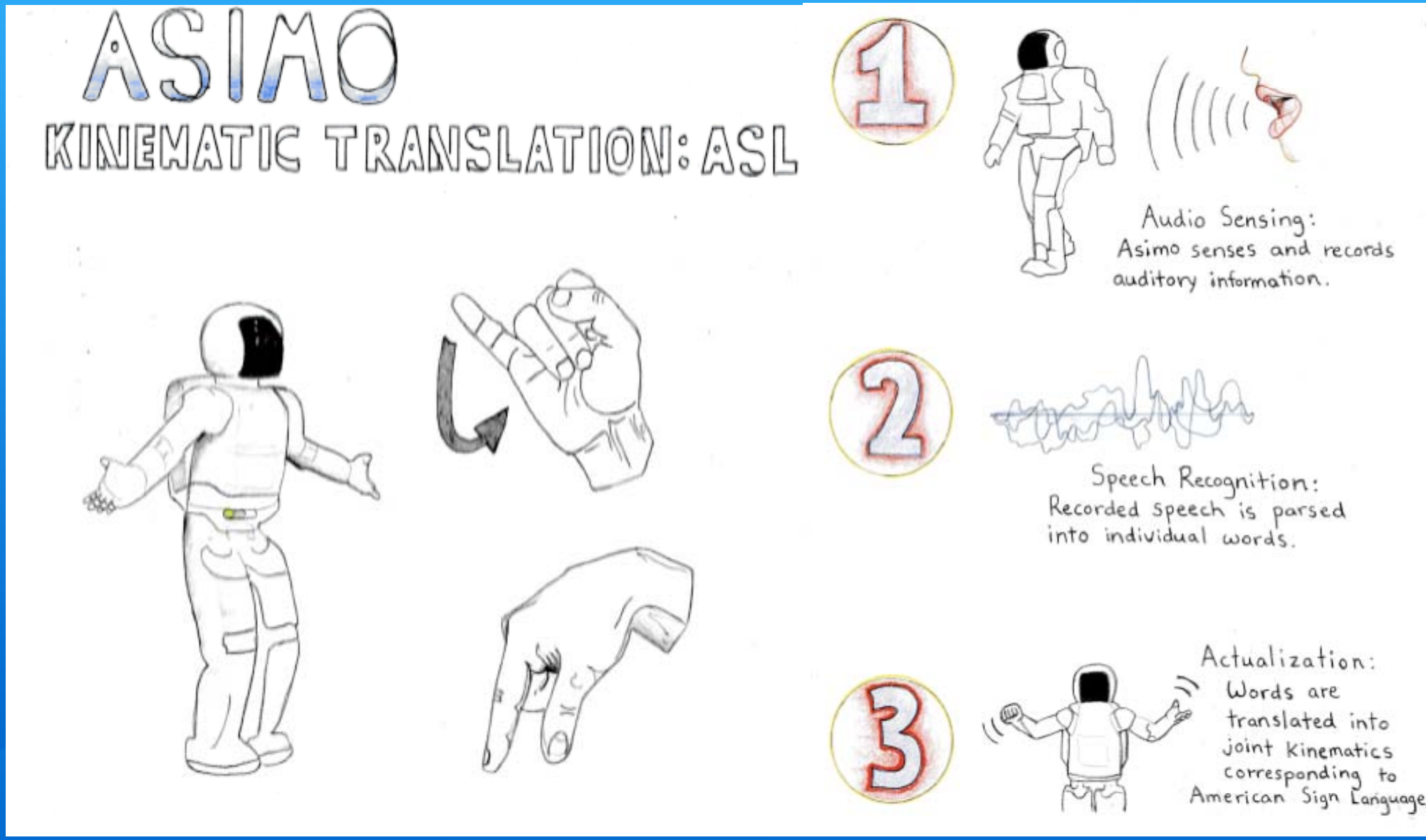
Guess these Humanoids!



Summer Project

Make Asimo Dance!!

The Concept


ASIMO
KINEMATIC TRANSLATION: ASL
1
Audio Sensing:
Asimo senses and records auditory information.
2
Speech Recognition:
Recorded speech is parsed into individual words.
3
Actualization:
Words are translated into joint kinematics corresponding to American Sign Language

Previous Efforts

- Nandy, A.; Mondai, S.; Prasad J.S. et al. *"Recognizing & Interpreting Indian Sign Language Gesture for Human Robot Interaction."* ICCCT, Sept. 2010

Image processing and genetic algorithms for feature extraction. Improves on existing HRI. Software developed for recognizing and generating ISL.

- Chiang Mai Univeristy/CT Asia Robotics: "Dinso" Thai Sign Language Interpreter. June, 2010

Visually processes human motion (Thai sign language) and translates into spoken sentences.

- Keita Matsuo, Hirotsugu Sakai (Fukuoka-ken Education Cetner): Japanese syllabic translator. 2006

80 cm 18 DOF robotic hand recognizes and translates the 50-character hiragana syllabary and about 10 simple phrases. Microchip/Servo-controllers.

Fig. 5: Japanese finger alphabet

Clockwise from top-left:

- "Dinso" Chiang Mai University
- Fukuoka Education Center
- Colorado State University

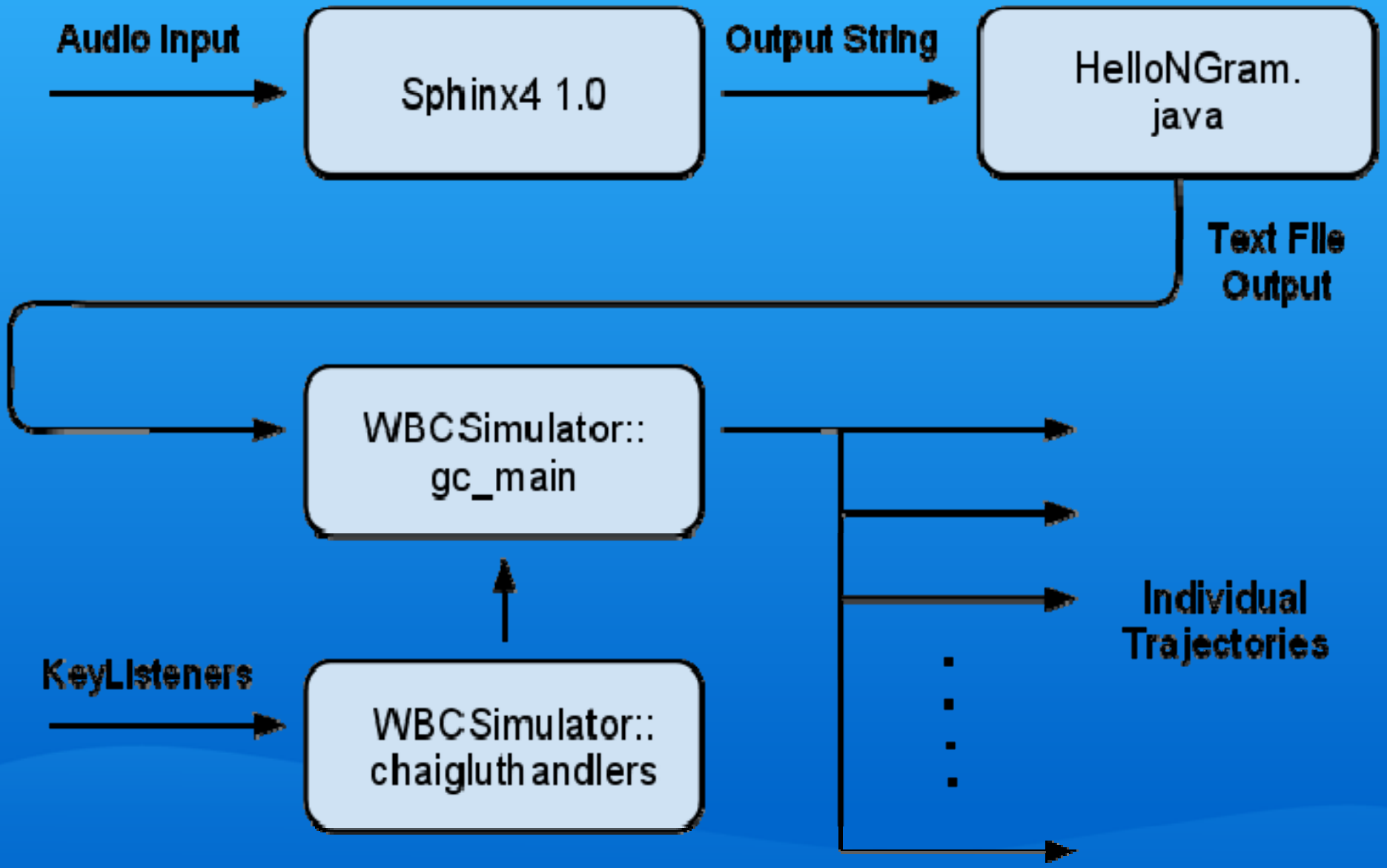
ASL: Background

- Many cultures have used non-verbal forms of communication for thousands of years (5th Century BCE- Socrates refers indirectly to system of signs used by the deaf).
- American Sign Language (ASL- formerly Ameslan) is related to Old French Sign Language (OFSL- a system of signs used by the deaf community in Paris in the mid 18th century). Brought to North America by Protestant minister Thomas Hopkins Gallaudet in 1815.
- ASL is a "natural language" containing phonology, morphology, semantics, syntax and pragmatics. Not related to spoken English.
- Manual nature allows for information to be loaded simultaneously on distinct channels of communication.

Actualization

- WBC Robot Simulator: Samir Menon, Stanford University 2010-11.
- CMU Sphinx Voice Recognition OSS.
- Sphinx processes auditory information- uses a learning algorithm to select appropriate English word from pre-defined lexicon.
- Strings parsed into individual character tokens and sent to simulator.
- Each character's corresponding representation in the ASL alphabet is signed in succession.

Program Architecture



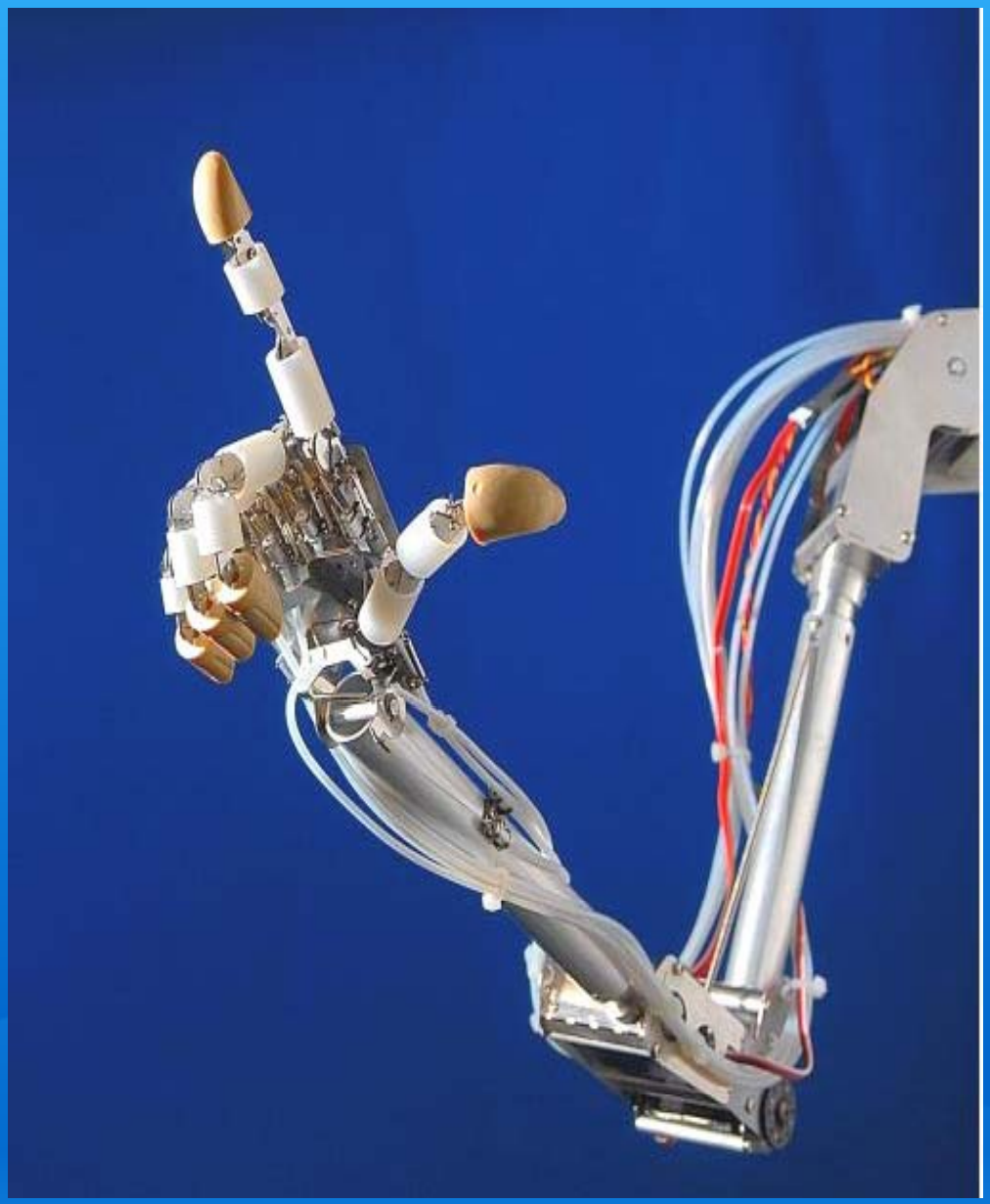
Trajectory Generation

- Joint-space control
- One trajectory for each alphabetic character

```
//resets hand+fingers
if (tstamp > 1) {
    sleep(1);
    tstamp = tstamp - 1;
    for (unsigned int i = 13; i < 24; i++) {
        gc_ctrl_ds->des_q_(i) = rob_io_ds->sensors_.q_(i)*cos(3.14*tstamp
    }
    gc_ctrl_ds->des_q_(12) = rob_io_ds->sensors_.q_(12) + tstamp*(-1.35 -
    if (tstamp > 1) traj_achieved = true;
}

if(true==traj_achieved) {
    wbc_chai_glut_interface::OPTION_RUN_SIM_K=false;
    traj_achieved=false;
    std::cout<<"\n End Sign K "<<std::flush;
    start_t=true;
```

Limitations/Future Considerations

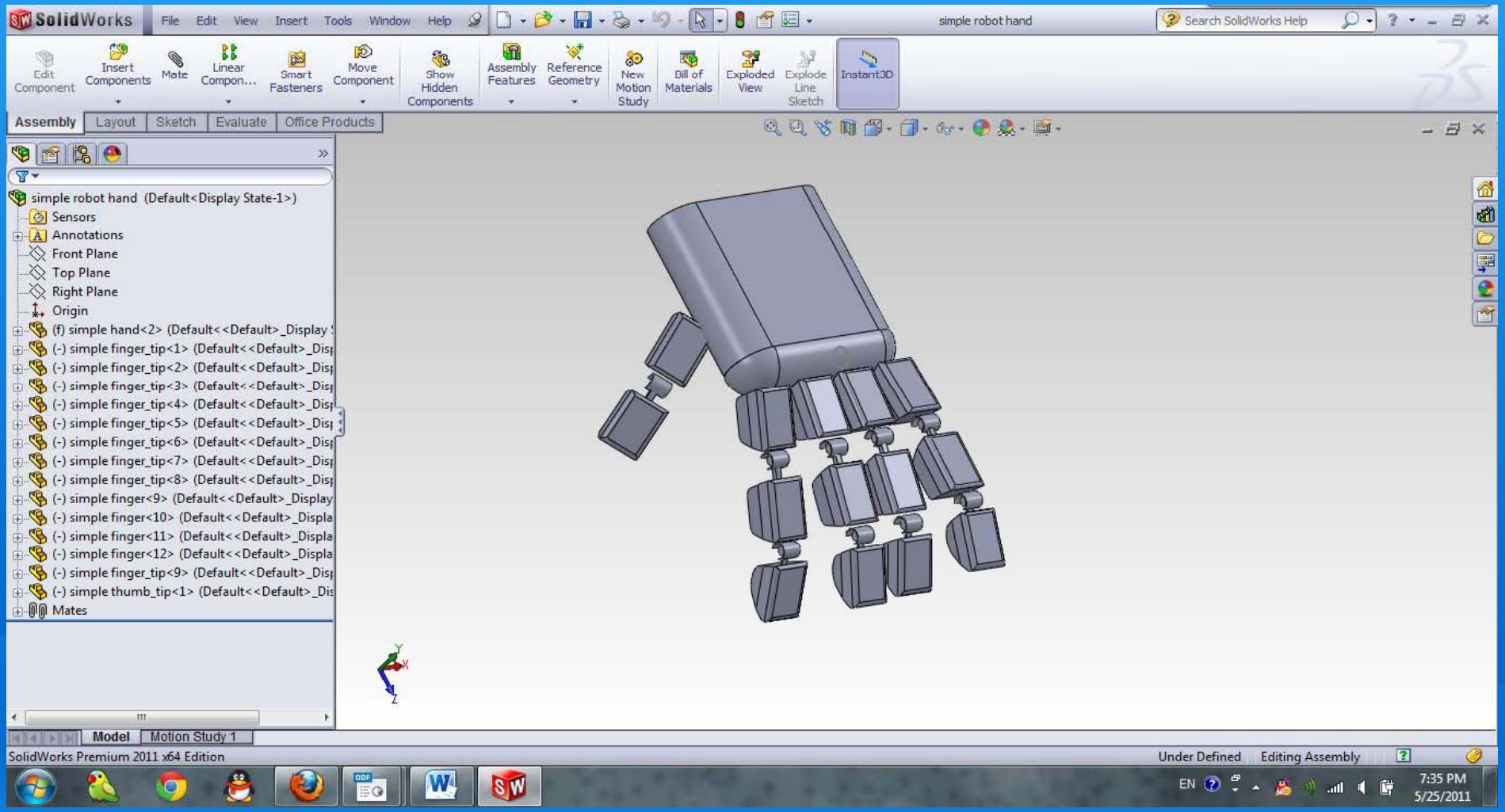


- Parsing ASL grammar
- Vocabulary
- Joint-space control vs. operational-space control
- Dynamic response (emotions, facial expressions)
- Actuator speed

$$\Lambda_{\otimes|s}\dot{\vartheta}_{\otimes} + \mu_{\otimes|s} + p_{\otimes|s} = \overline{J}^{T}_{\otimes|s} N_s^T \Gamma$$

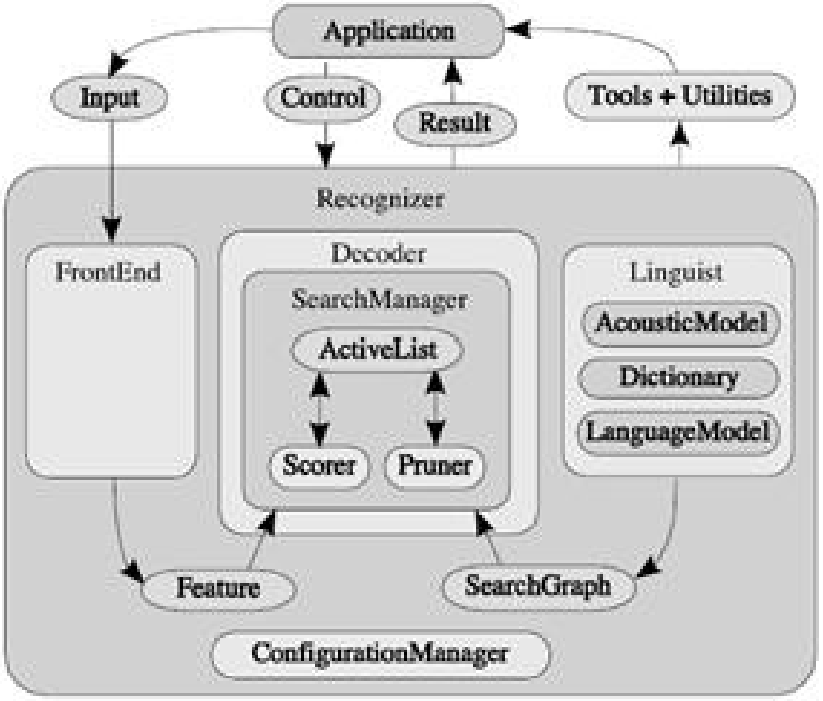
Graphics

D.O.F of the hand = 3 + 1 + 2*4 =12



Voice Recognition

- Sphinx4
- Hello N-Gram Demo




```
Start speaking. Press Ctrl-C to quit.

You said: the green one on the right side

Start speaking. Press Ctrl-C to quit.

You said: it's the uppermost of three in a row

Start speaking. Press Ctrl-C to quit.
```

Real-time Demo

Final Thoughts…