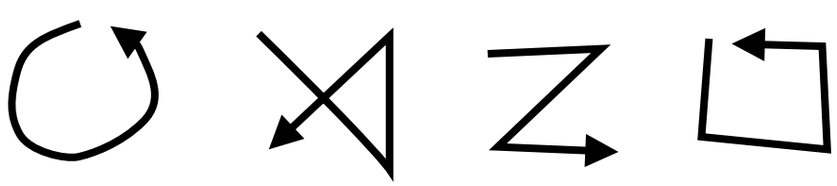
3D Gesture Recognition

Ben Alun-Jones

Gesture Definition

- Loop, X, Zig-Zag, Box : Mode selection



- Shake : Discard, undo

- Can extend to more gestures later

Learning Gestures

- Gestures:
  - are time varying, location, socially varying
  - denote meaning and governed by rules
- Similar to speech but:
  - are complex - sequence of actions for one gesture

Learning Gestures

- Need to use machine learning

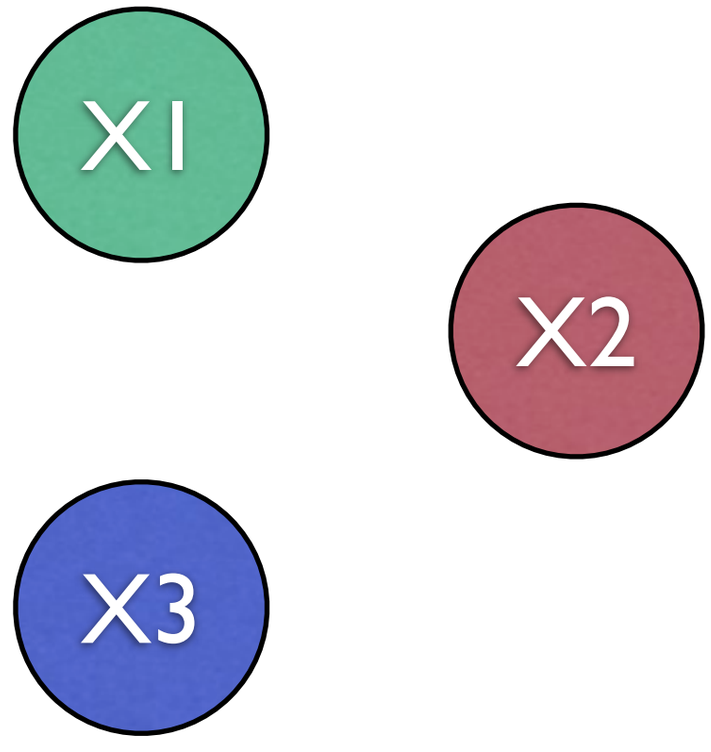
> *Machine learning is a scientific discipline that is concerned with the design and development of algorithms that allow computers to evolve behaviours based on empirical data, such as from sensor data*

- Hidden Markov Model is an example of such a technique

Markov System : Example

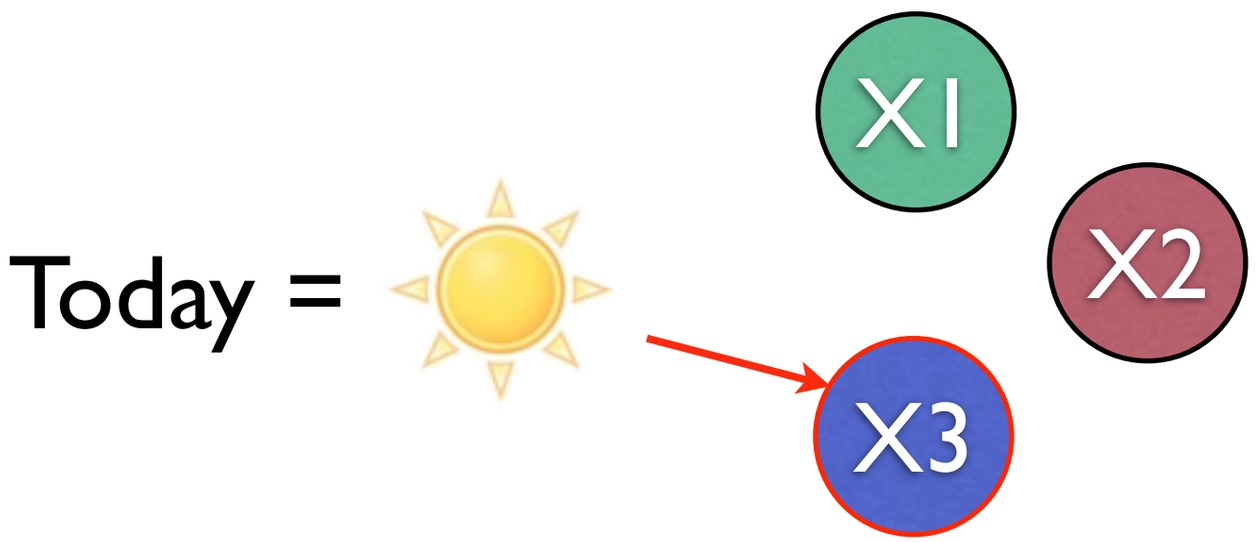
- Has N states (3 in this example)
- Assume discrete timesteps (t = 0, t = 1, ...)

e.g. weather



Markov System : Example

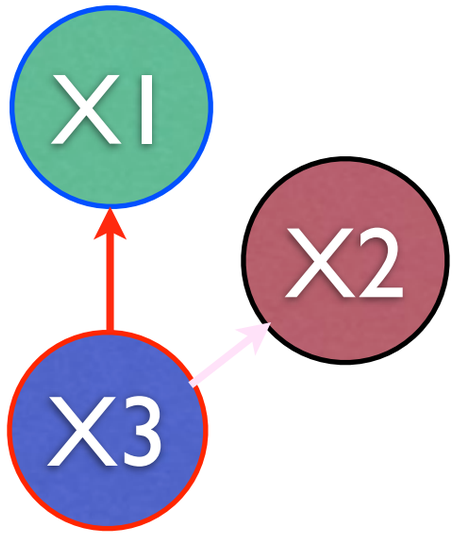
- At time t=n, the system is at exactly **one** of these states



Markov System : Example

- The next state is random
- Determined exclusively by today's weather **not** how we got to today.

Tomorrow = 

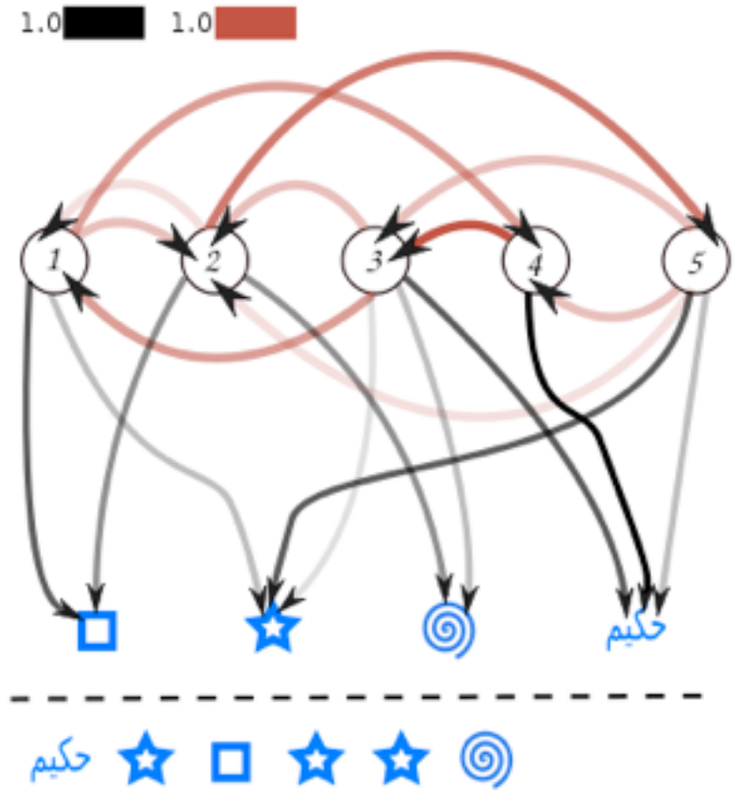


Markov Property

- The Markov property states that the probability distribution for the system at the next step (and in fact at all future steps) only depends on the current state of the system, and not additionally on the state of the system at previous steps.
- Google Page Rank is a Markov Chain over the internet

Hidden Markov Models

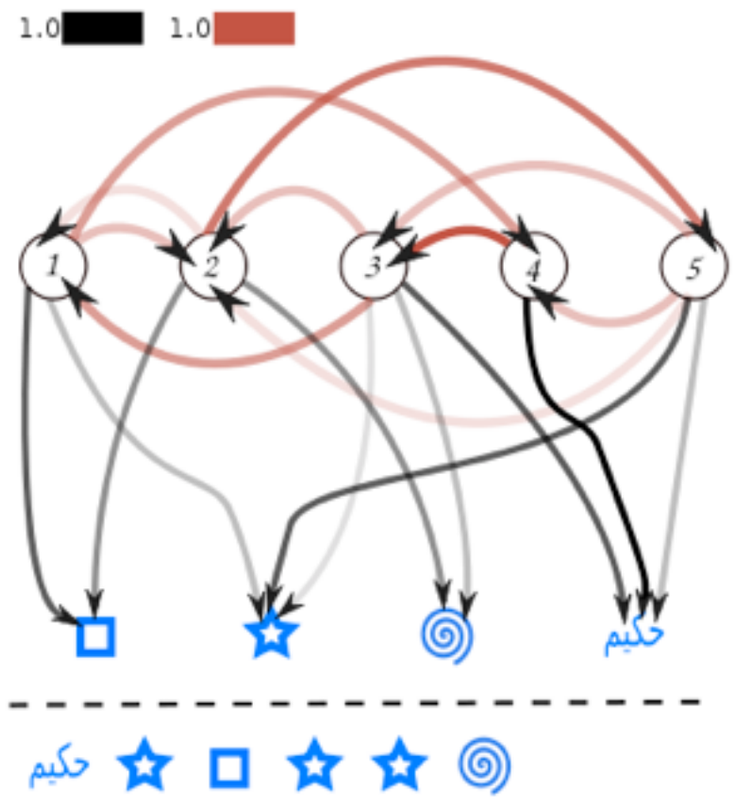
- Characterised by:
  - A set of states
  - Transition probabilities
  - output probabilities



- e.g. determining the weather by speaking to your friend on the telephone

Hidden Markov Models

- Cannot see the state we are at, only the output of that state.
- We observe **only** the outputs
- The ‘Learning’ part is given these outputs, can we determine the underlying states and their transitions



Hidden Markov Models and Gesture recognition

- Have the Observation Data
- Sequence of actions determines the gesture
- If the sequence observed, can then say it was a particular gesture

Hidden Markov Models and Gesture recognition

- Can solve three basic problems:

1. **Learning** : Provide some model parameters so model has high probability of generating observations for given data

Hidden Markov Models and Gesture recognition

2. **Evaluation** : Score the matches between the model and the sequence.

Isolated gesture recognition

Hidden Markov Models and Gesture recognition

3. **Decoding** : Finding the best state sequence for a given observation sequence.

Continuous Gesture recognition

Methodology

1. Describe each gesture
2. Collect training data
3. Train the HMM
4. Filtering

1. Describe each gesture

- 3D acceleration, 3D velocity, 3D Position
- 9D observation vector possible
- 5 unique gestures specified
- Transition Matrix and output matrix values determined by training

2. Collect Training Data

- Because of Markov property, each data ‘point’ can be processed independently
- Need to pre-process e.g. vector quantisation, mapping 3D to a 2D plane, Fourier transform, PCA
- Vector Quantisation by k-means shown to work well

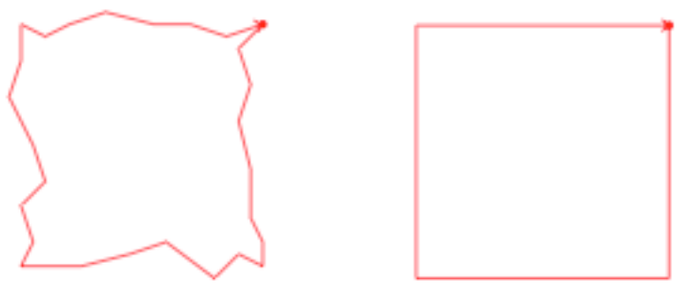
4. Train the HMM

- No analytic solution
- Simply maximise the likelihood (the probability of correct observations given the model)
    - Iterate the model using the Bauch-Welch Algorithm to add
    - Classify gestures using left-to-right algorithm

5. Filtering

- Crucial stage to make for robust recognition

1. Filtering of meaningless vectors:



(a) Before filtering (b) After filtering

2. Determine when the accelerometer is idle

Future Steps

- N900 difficult to program but close
- Wii Remote possible
- Want to make a comparison of gestural control versus current HCI paradigm e.g. in the Allosphere

Conclusion

- Using Hidden Markov Models can determine underly truth of sequence from observations
- Unable to demo as progress limited by hardware