Negotiated Interaction

Iterative Inference and Feedback of Intention in HCI

Roderick Murray-Smith, John Williamson

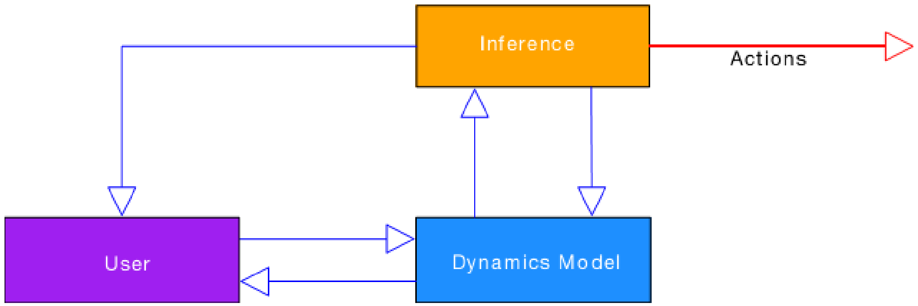
Department of Computing Science,
University of Glasgow &
Hamilton Institute, NUI Maynooth
rod@dcs.gla.ac.uk
http://www.dcs.gla.ac.uk/~rod
http://www.dcs.gla.ac.uk/~rod/Videos.html

Bayesian Research Kitchen, Grasmere, 7th Sept 2008.

Hamilton Institute

Negotiated interaction

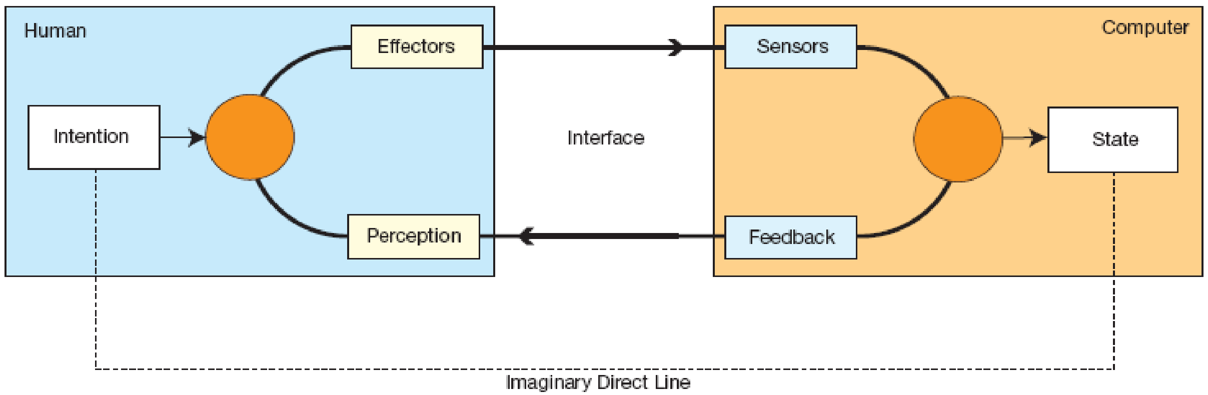
- A new framework for interaction design could include:
  - Users interact with content, services and other users in environment
  - Actions and feedback can be continuous
  - User and system negotiate interactions and intentions in a fluid, dynamic manner.
  - Dancing metaphor, rather than command-and-control. Ebb and flow of control, changing fluidly as context determines.
- Sharing the load
  - The interaction problem viewed as a negotiated control process, where user and system work together to communicate intention.
  - Timed, informative feedback shares the load between both sides.
  - This occurs at multiple time-scales



My perspective on Interface Dynamics

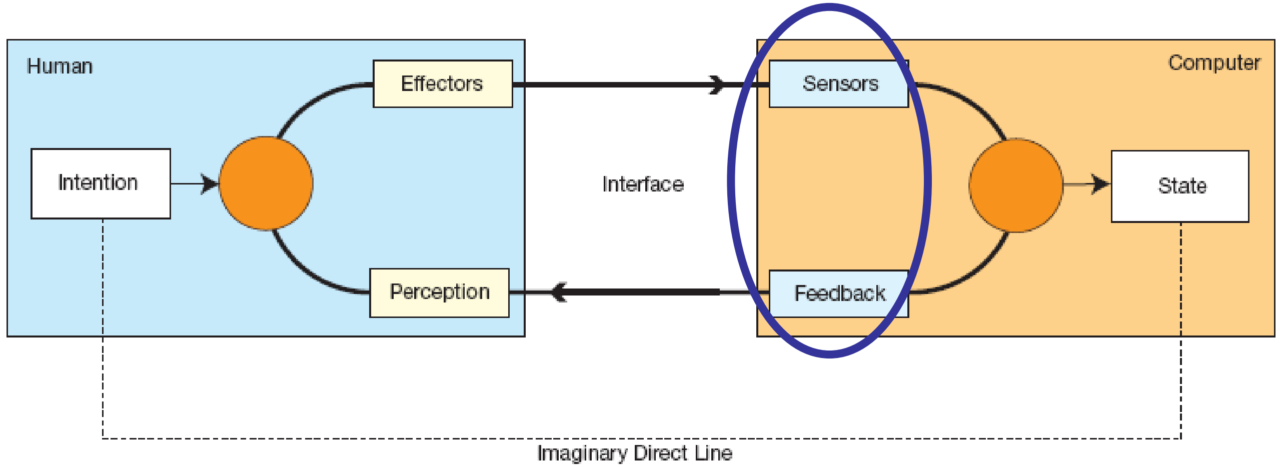
- Control theory perspective
  - We have evolved to control our perceptions. We require feedback, and there are upper limits on our bandwidth.
  - User interacting with interface object viewed as two coupled dynamic systems
  - Physical model-based approach to representation of interface objects
  - Dynamics allows us to slip in ‘intelligence’ into the closed-loop which couldn’t be done with a static interaction technique
- Probabilistic perspective – uncertain interaction
  - Uncertainty in user’s mind about what to do next, and system uncertain about user’s intentions.
  - Dynamics and feedback are adapted based on probabilistic inference.
  - Taking explicitly Bayesian view. Probability distributions will be assigned to beliefs in a system.
  - Joint system dynamics mediate the flow of evidence between participants at an appropriate rate.
- Multimodal, embodied perspective
  - Coupling and interaction is continuous (time and space) and feedback is multimodal.
  - Interaction is active – energy in, information out.

Interaction as closed-loop design



- The interface is a mechanism for controlling the flow of information from a system
  - an interactive system has therefore to ascertain the intention of the user with the minimal effort on the part of the user.
- The interaction is formulated as a continuous control process, where the system is constantly engaged in recursively updating a distribution (inference) over the potential intentions of a user while providing feedback of the results back at a range of timescales, which users can then compare with their goals.
- User and system attempt to negotiate a satisfactory interpretation of the user's intention.

Novel sensors and displays



- Wide variety of sensing and display technologies that can be used to construct the physical aspects of a human-computer interface.
  - Rich sensors, from accelerometers, to smart clothing, to GPS units, to pressure sensors etc, create the potential for whole new ways of interacting with computational devices in a range of contexts.
  - Each of these has different information capacities, noise properties, delays, frequency responses, and other modality-specific characteristics.
  - Sensors will get cheaper, and new ones will create as yet unimagined interaction possibilities
- Building interfaces that make use of possibly *high-dimensional, noisy*, *intermittently available* senses to create usable communication media is a challenge.
- We need general frameworks which are not tied to specific sensing or display devices, but generalise to wider classes of devices.

Midas touch

- How do we control the interpretation of our phone’s sensor readings? How do we ‘declutch’ certain modes?
- Sensor flow will be interpreted differently in different contexts
- Needs excellent models to automatically infer likely intention given overt behaviour.
- Need subtle feedback to user for them to infer current mode & consequences of action.
- This is a major, fundamental area which will recur everywhere in mobile multimodal interaction.

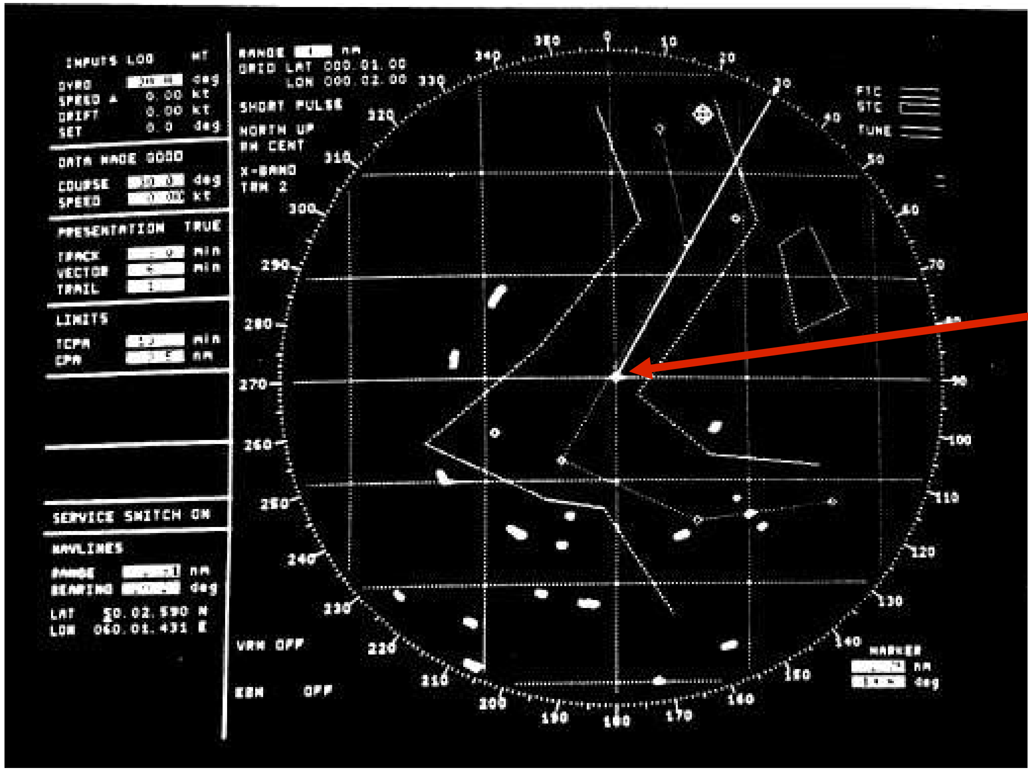
Feedback Modes

The display is to provide the user with information needed to exercise control. i.e. predict consequences of control alternatives, evaluate status and plan control actions, or better understand consequences of recent actions.

- Basic feedback loops
  - Visual, audio, vibrotactile display of states of phone, or of distant events, people or systems.
- Modality scheduling
  - Order of presentation of information in different feedback channels.
- Mobile context
  - Disturbances, lower attention span, fragmentary/intermittent interaction.

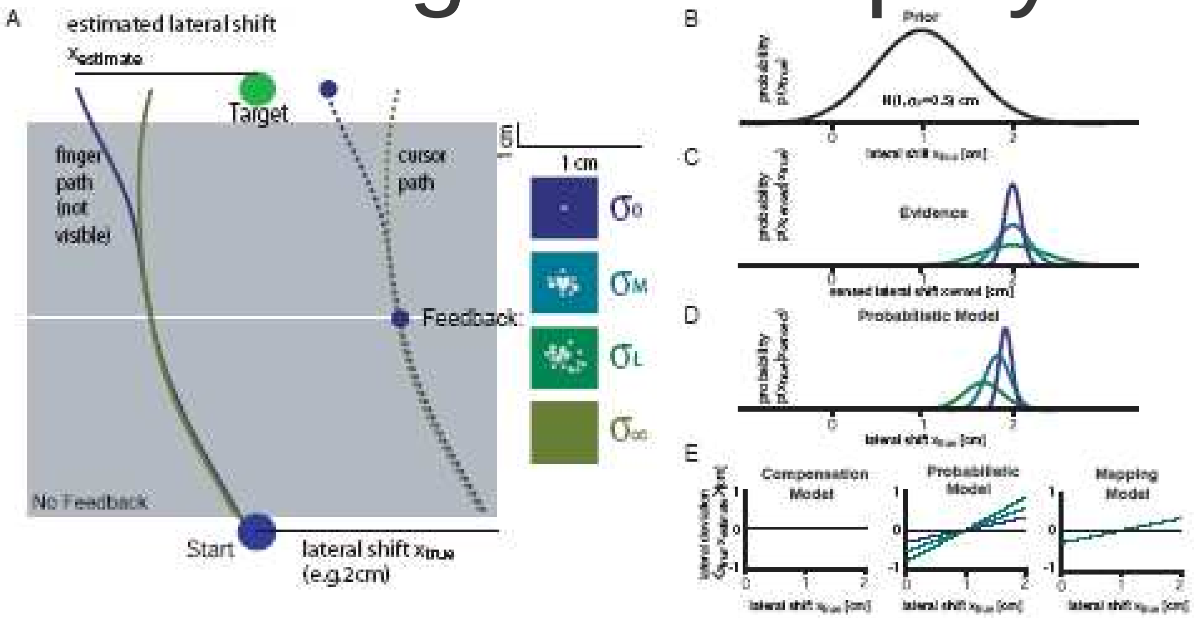
Uncertain Display

- Poor displays lead to poor control
  - Classic example of The Royal Majesty



"precise" position

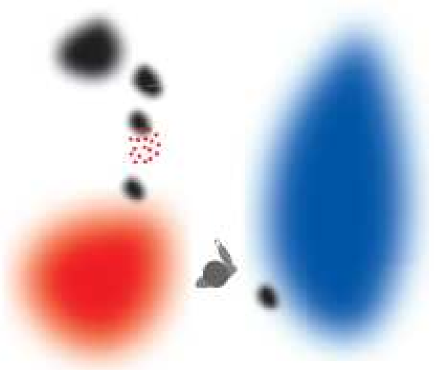
Ambiguous displays

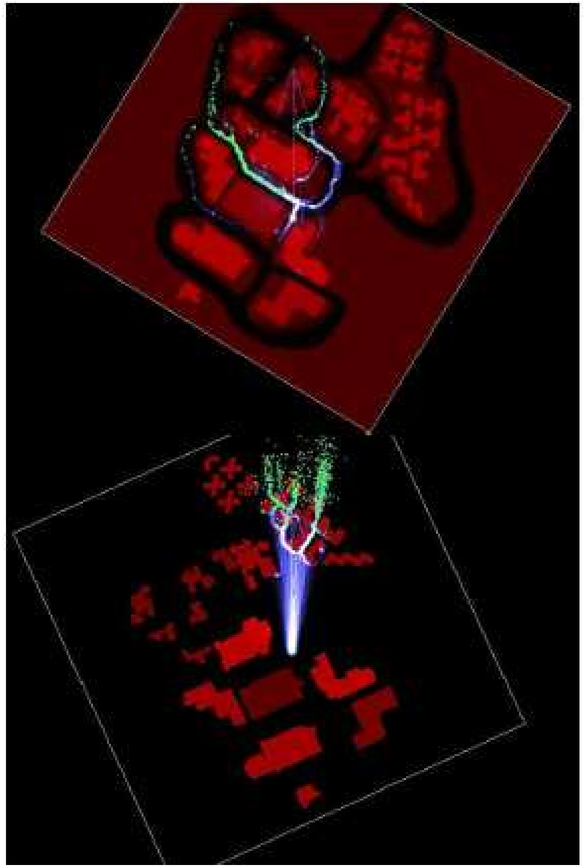


- Used in psychophysics experiments (e.g Körding & Wolpert 2004)
- Transfer idea to user interface design. If the system is uncertain about inputs or user intentions, present data in an appropriately ambiguous fashion.
- Does it regularise user behaviour & improve usability appropriately?
- Pattern recognition and displays are interdependent and should be developed together

Particle GPS Browsing

- Location-aware audio & haptic feedback
- Use tilt and bearing to get rapid exploration
  - Project forward, find likely locations in the future.
- Map browsing; include uncertainty about where we are
  - Show all the possible places we might be, given a map of the area
  - User can scan around and project further into the future.
- Augmented reality content is interpreted by models which generate multimodal feedback



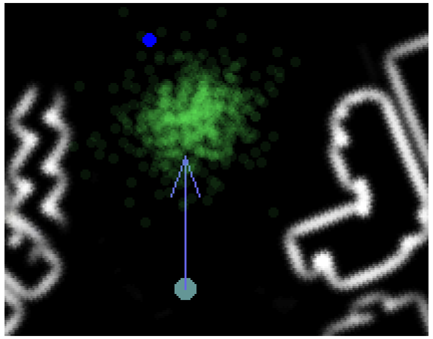


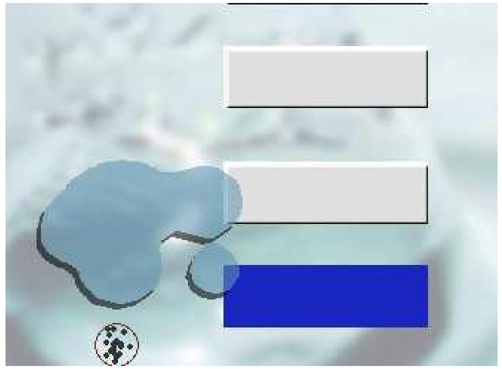
Liquid representation of interaction

Spreading inference over time

- Belief state of system is high-dimensional
- How can we drive it to a particular state?
- Human actions are noisy, imperfectly controlled, and imperfectly planned. Interface sensors measure activity in non-transparent ways
- Mapping from user intended communication and what is measured by system's sensors is a complex, uncertain mapping.
- Real-world interaction always involves control
  - People receive feedback about the consequences of their actions
  - By breaking down the task into a physical control problem inference of intention can be spread out over time, and the limitations of human action and computer sensing systems can be overcome.

Liquid, gas, solid…

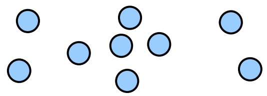




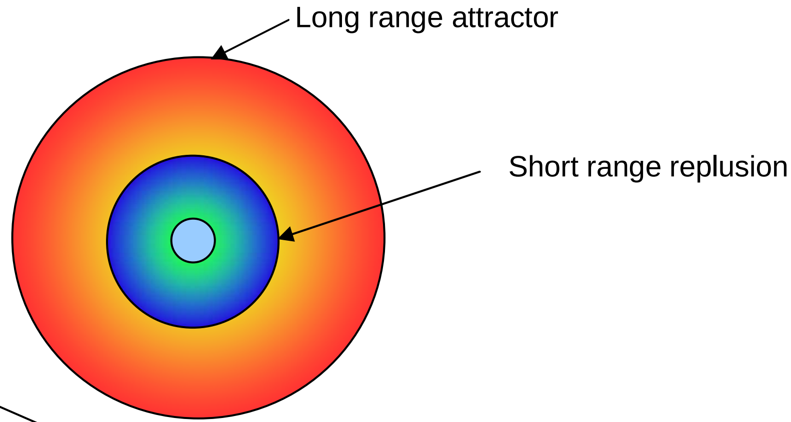
- Gas (MC) shows inferred beliefs, but is less focussed on action and control
- Solid point has no distribution, therefore limited feedback for user. Has clear control only when using low-noise, directly mapped inputs.
- Liquid form is not a true distribution, but does relate to control, and is better suited for guiding the user's attention.
- Potential for dynamic change of properties (moving from true distribution to negotiated one?)

Liquid Cursor

Start with Monte Carlo samples

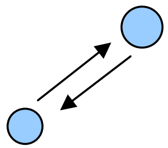


Equilibrium of attraction and repulsion
(with damping)



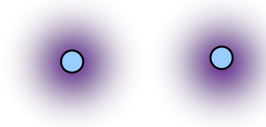
Add molecular dynamics

Particles exert force
on each other

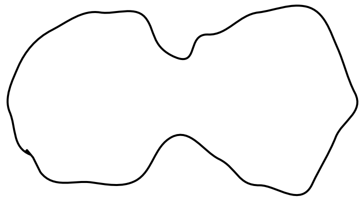


Render with isocontour tracing

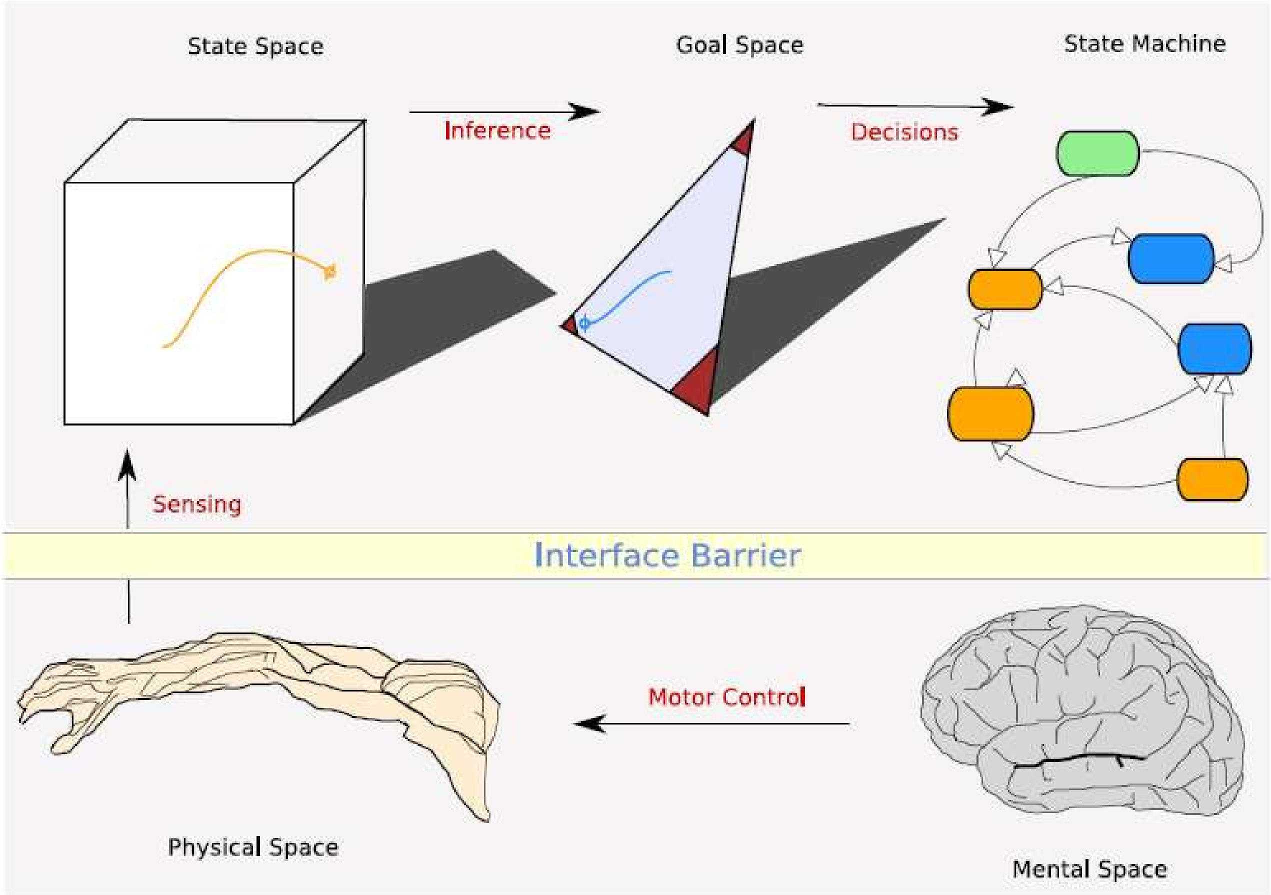
Gaussian on each sample



Render the isocontour

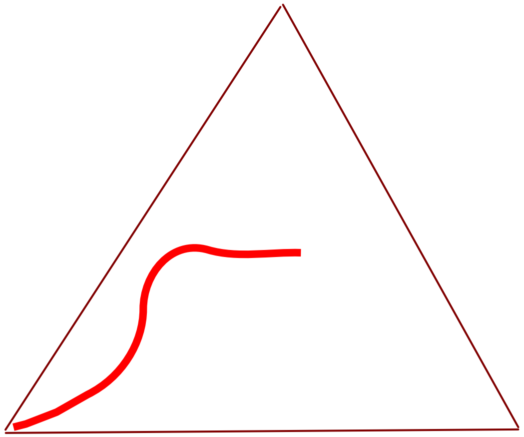


Evidence, Goal and State spaces



Goal Spaces

- We focus on the problem of interaction with sensors producing continuously varying measurements.
- The interaction is a closed-loop control process and the ultimate control variable is the distribution over actionable goals.
- The purpose of the system is to perform recursive evidence updates to infer the new goal distribution, forming a trajectory through the space of distributions. The space in which this trajectory lies is the goal space;
- For example, discrete selection: $p_1 \ldots p_n$ simplex in n-d space
  - Inference (should) result in a smooth trajectory in this space
  - Large steps in entropy are unnatural & error-prone
  - Information rate determines smoothness
- Give feedback to user about progress through this space. By avoiding discrete state changes as long as possible, the need for after-the fact correction system such as *undo* can be minimised.



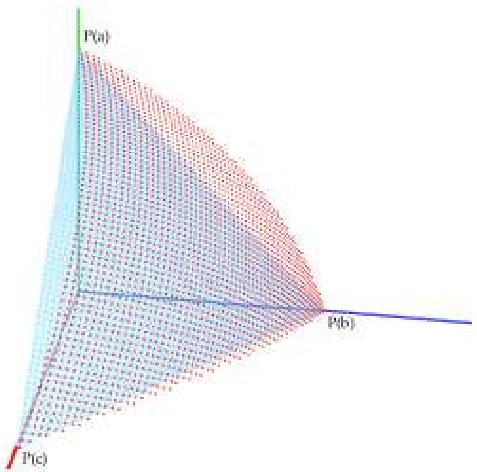


FIGURE II.5: The changing entropy H(x) across regions of the goal space, shown here for a three goal system. Entropy is shown as the dotted surface above the simplex – distance from the surface indicates entropy at that point. It reaches its maximum of 1.584 bits at $< \frac{1}{3}, \frac{1}{3}, \frac{1}{3} >$, and drops to zero at the vertices.

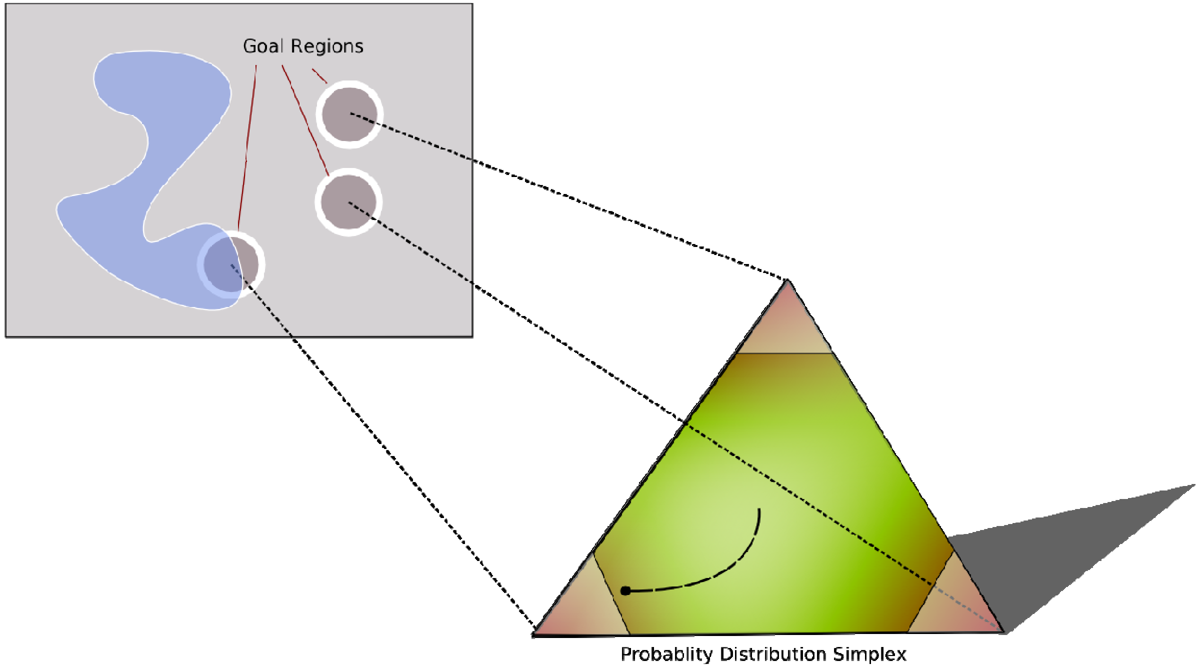
Information and Smoothness Constraints

- If a point x in the goal space is considered, $H(x) = -\Sigma_n p_i \log_2 p_i$ is the Entropy at that point. The communication rate of the system is given by $dH(x)/dt$.
- There is assumed to be a maximum potential communication bit-rate b_{max} – the information capacity of the interacting muscle group is one such upper bound, for example; the sampling rate of a sensor is another.
- If the process is to be controlled by the interactor, however, the bandwidth of the feedback must also lie within the user's ability, as otherwise the interaction will be unpredictably unstable.
- So $b_{max} = \min(b_{maxin}, b_{maxout})$. b_{max} enforces a smoothness constraint on the goal space trajectories; since $dH(x)/dt \leq b_{max}$.

Maximum Information Limit: Prohibiting Excessive Bandwidth

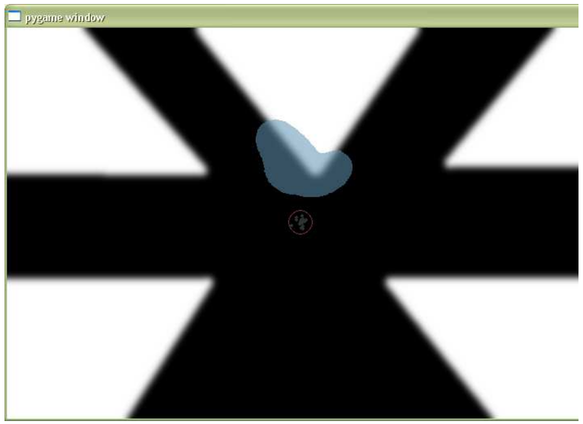
- Well-designed systems should have smooth trajectories in the goal space
  - large jumps indicate either that:
    - evidence has been too slowly sampled (e.g. in a keyboard system, where only the terminal result is available as a discrete decision, although this will still obey the bit-rate law *on average)*.
    - little feedback can have been provided, or that excessive weight is placed on evidence and decisions are made without basis.

Link between display and goal spaces



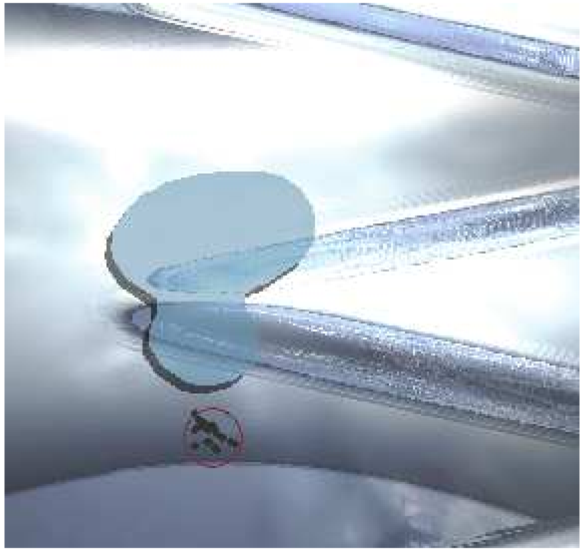
- Liquid cursor is 2-D as in existing pointing techniques
- Dynamic properties allow gestures to be recognised.
- Multiple hypotheses can be maintained until sufficient evidence is provided to effect an action
- System and user share a model of the distribution over targets

Adapting the fluid dynamics



- Liquid viscosity can be varied according to derivative of entropy of intention interpretations.
- Have multi-component liquids with different viscosities associated with different time-scales.

- Liquid cursor acts as coordinating medium
  - Multiple sources of evidence are combined in real-time in a visually obvious manner
  - Updates of evidence have immediately tangible effect on the form of the liquid
  - Prior beliefs can affect the flow of the liquid, essentially creating attractors around likely beliefs and repulsing constraints around unlikely ones.




pygame window


pygame window
?

Cromwell’s dictum & Undo

- space of potential states of the system explodes exponentially
  - the external world must be affected at some point.
  - The number of decisions that can be kept reversible has a significant effect on the usability of the interface.
- Undo is necessary for three reasons:
  1. A user was unable to predict the response of the system and so performed the wrong action;
  2. A user attempted to, but was unable to perform the appropriate action (for example because of physical slippage);
  3. the user changed intentions (e.g. the user was exploring the capabilities of the system, and decided that the action performed was not the appropriate one in retrospect).

Semantic Pointing (Blanch, Guiard, Beaudouin-Lafon 2004.)

- Motor space and Display space have different properties
- Control-Display ratio adapted depending on proximity of target

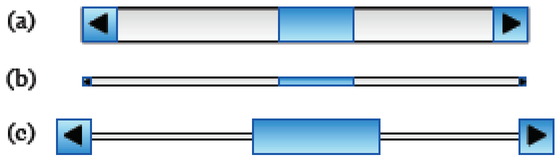


Figure 12: Scroll-bar redesign
(a) original version. (b) new version: visual space (what it looks like) and (c) motor space (what it feels like when interacting with it).

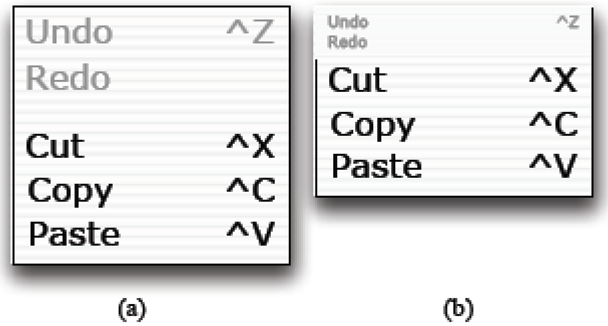
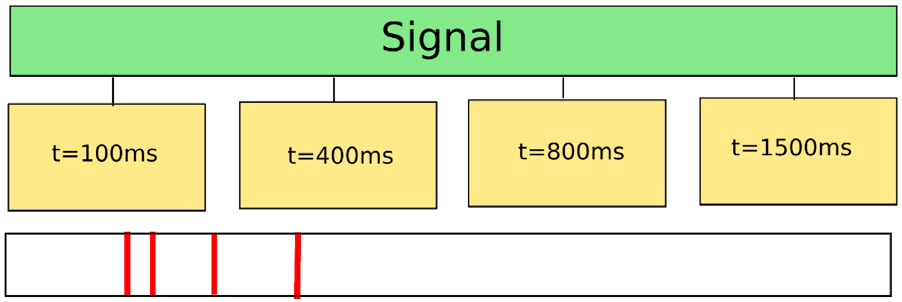


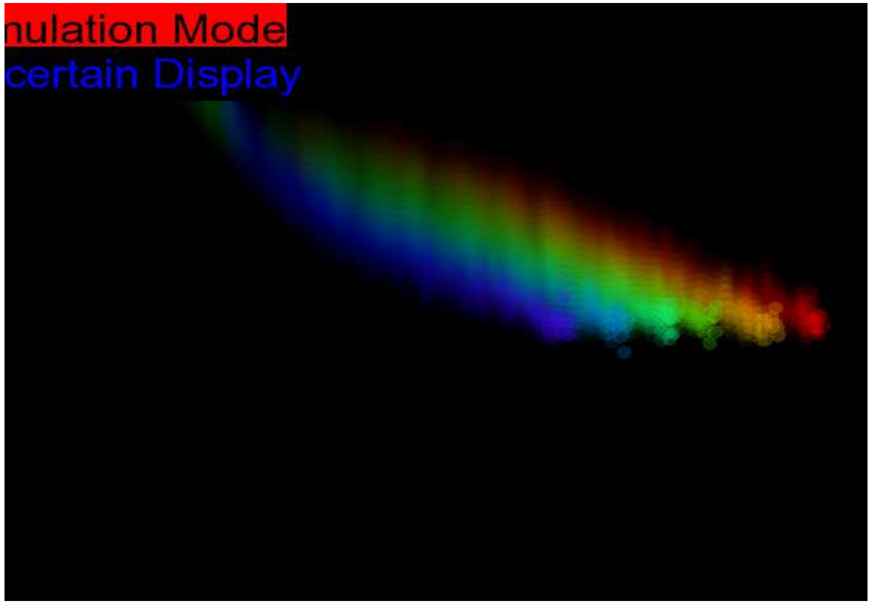
Figure 13: Menu redesign
(a) unchanged visual version (b) motor space version

Figure 14: Button redesign
(a) unchanged visual version (b) motor space version

Uncertain Multiscale Multimodal Feedback in BCI

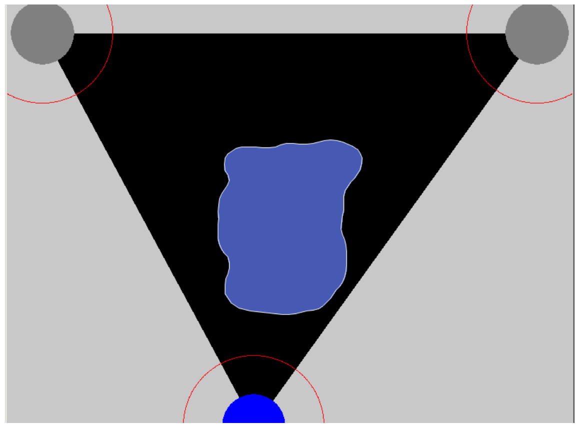
- Each timescale represented visually
- Point “cloud” to represent uncertainty

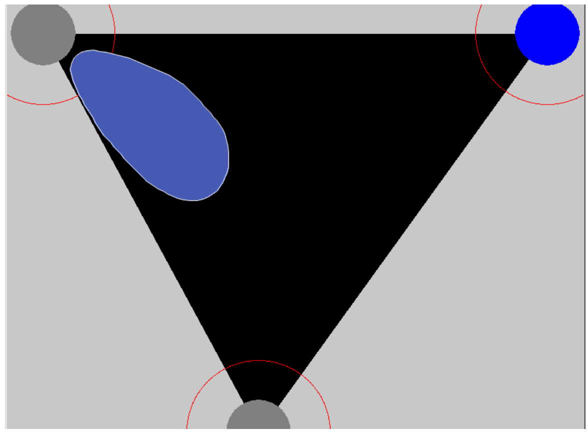




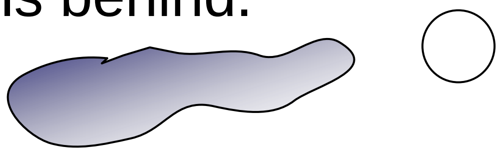
Multi-Class Liquid

- Instead of point cloud, create **liquid simulation**
- Move on space of potential possibilities
  - Goals at corners
- Dynamics are revealed by the blob's shape changes

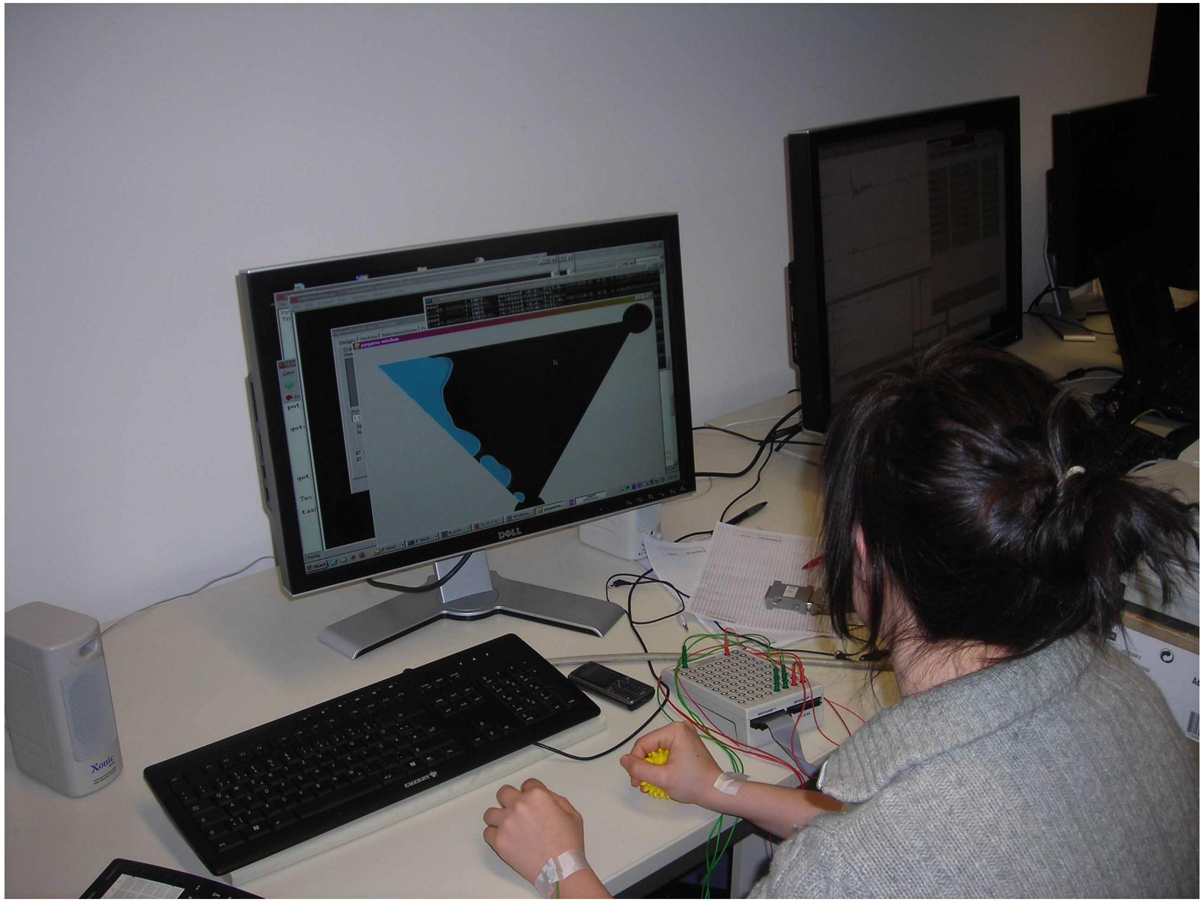




Could also do multi-timescale, with blobs with excitable heads heaving tails behind.

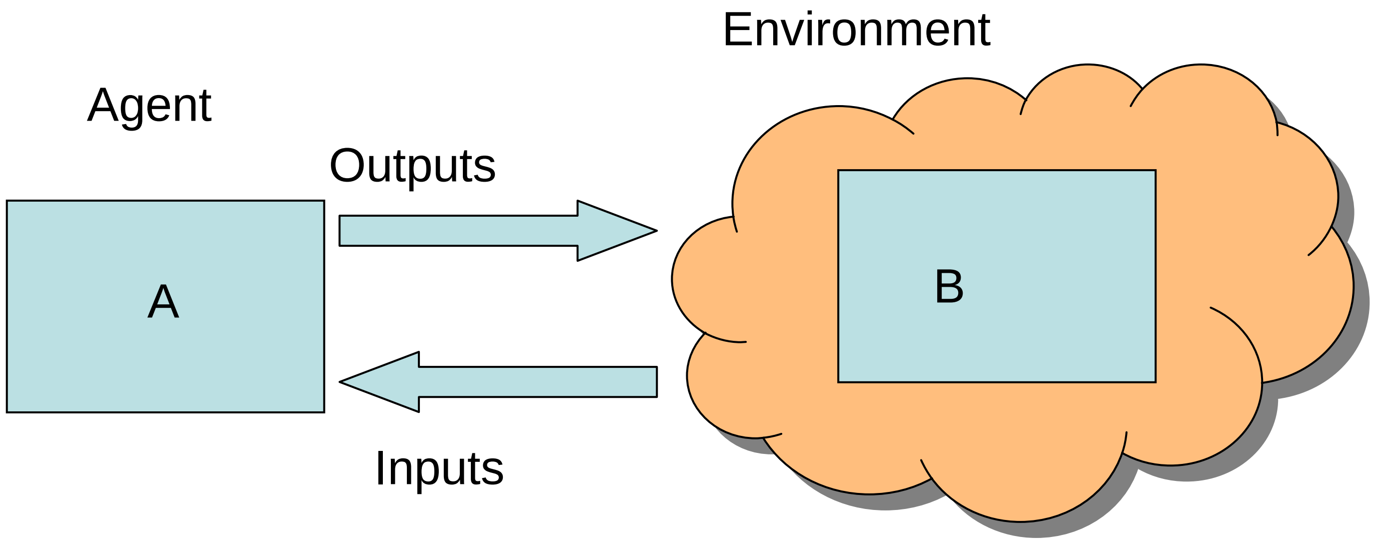


Testing with EMG input



Measuring Interaction?

Empowerment – interaction as control



- Empowerment is the maximum flow the agent can direct into its future sensoric input via the environment
  - "All else being equal – keep your options open". Striving for more options, with more potential for control or influence.
- Measure of control suggested by (Klyubin, Polani & Nehaniv 2006), building on work of (Powers 1972).
  - Information-theoretic capacity of an agent's actuation channel.
  - Channel capacity is the maximum mutual information over all possible distributions. It is asymmetric and causal, and requires control over X.
  - How directly is output from agent B going through A and back to B?
- How does a control perspective change how we think about design?

Measuring Interaction

- Interaction design is of great importance, but little work on definition of measures of interaction.
- Many HCI textbooks do not explicitly define interaction. An example definition, *“By interaction we mean any communication between a user and a computer, be it direct or indirect”* [Dix, *et al.* 2004] does not provide an obvious way to measure the communication.
- We also need more detailed definitions which can take into account which elements of the communication actually make a difference.
- Why bother?
  - It could be the foundation of a more consistent framework for the study of HCI.
  - Measures of interaction in specific trials could augment subjective measures in usability studies
  - Adaptive, learning interfaces could use it as a cost-function to be optimised.

Developing a measure

- Any measure chosen will implicitly or explicitly incorporate a model of human behaviour.
  - Challenging, but already standard for low-level processes.
  - Key issue is that our framework should be able to cope with model uncertainty
- The more uncertain the models are, the less powerful the measure will be in any specific exchange between human and machine,
  - but it might still provide the optimal approach to designing an interface, given our uncertainty about human behaviour.

General definitions of Interaction

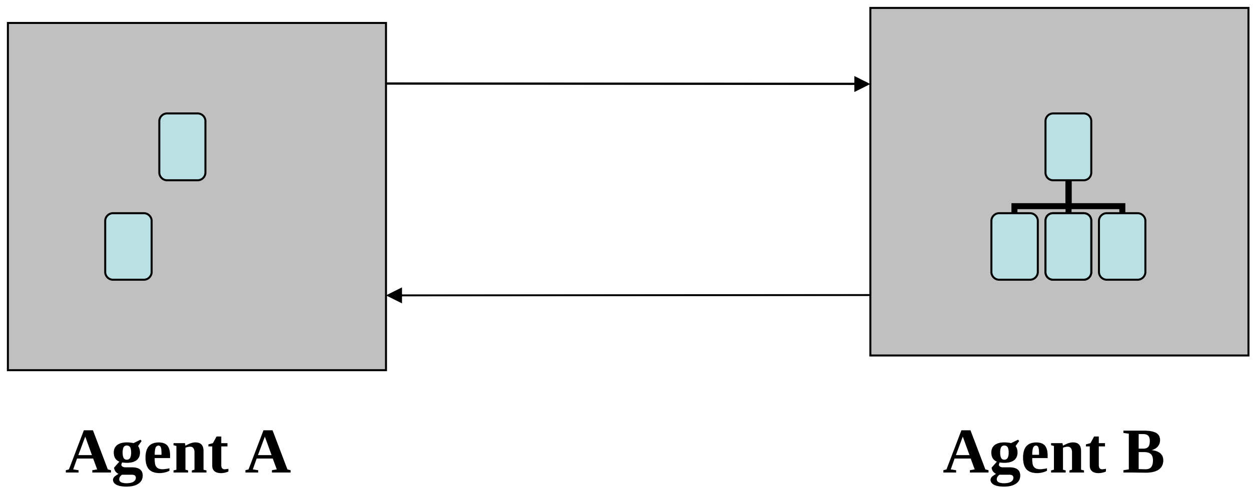
- Interaction is a kind of action which occurs as two bodies have an effect upon one another.
  - The notion of two-way effect is vital, as opposed to a one-way effect, where one system ‘drives’ the other.
  - Interaction occurs when humans and machines control each other’s behaviour (including the special case of communicating with each other).
  - It can occur whether the control and communication is intended or unintended.
- One definition is *Interactivity as degree to which an action is related to earlier actions between two agents*.
  - However, it is not clear that we should limit ourselves to past actions.
  - Most intelligent agents will be making predictions about future actions, and we can therefore have interaction occurring *before* the first action is made.

Possible measures

Three approaches:

1. Information theory
2. Predictive control/Game theory
3. Empowerment/control

Also strong links between causality measures in diverse fields and interaction measures



1. Information theory – Mutual Information

- Measure interaction in bits per act for discrete acts and bits/s for continuous.
  - In continuous case, need to integrate over different timescales
- I(X,Y) is a function of both the transmitted signal p(x) and the channel characteristic p(y|x).
  - I(X,Y) is symmetric in X,Y so is *acausal*.
  - We are more interested in causal measures – humans are acting as controllers.

2. Predictive control

- Use mutual predictions between agents
  - like dual control, the actions are trying to achieve a goal and probing at the same time.
- No general analytic solutions
  - look at inter-sensitivity between systems on actual interaction trajectories, via Monte Carlo simulation.
- If an agent is *engaging* with another, it can be said to be sensitive to changes in behaviour.
- Links to Game theory.

3. Empowerment

- Measure of control suggested by (Klyubin, Polani & Nehaniv).
  - Information-theoretic capacity of an agent's actuation channel.
  - Channel capacity is the maximum mutual information over all possible distributions. It is asymmetric and causal, and requires control over X.
  - How directly is output from agent B going through A and back to B?
- Qualitative observations:
  - If B has full understanding of A and controllability, then can generate its desired perceptions.
  - Unpredictability seems important.
    - If agent A is controllable and predictable, then no interaction - it is just an encoding problem.
    - If not fully predictable, then need to take feedback into account. Interaction!

Empowerment

- More of a focus on actuation, and naturally links perception and action
  - Not all actions lead to perceivably different results
- Timescale over which empowerment is calculated is a key issue
  - (relevant for battery life consequences of interaction?)
- How does making a change to your Facebook status entry link to empowerment?
- It is ‘interesting’ to be close to objects you can manipulate, as that increases the degrees of freedom.
  - Relevance for Mobile Spatial Interaction?

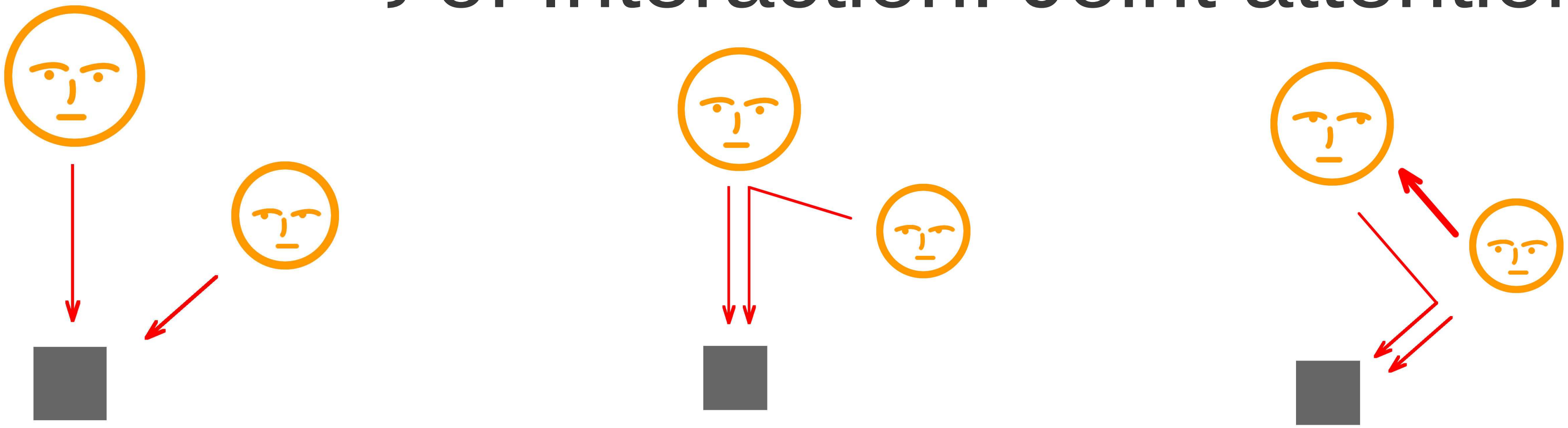
Predictions of interesting behaviour

- Interaction levels can increase *before the first actions occur* (effect of prediction)
- Decreased level of interaction as uncertainties and delays increase.
  - Uncertainties reduce the prediction horizon
  - Delays limit power of feedback to compensate for poor models
  - One important uncertainty is due to the effects of internal reference values of each agent which are hidden but can be inferred from actions.
  - Effect of initial conditions. Two agents might give quite different measures of interaction, depending on where they start.
- Would behavioural “bottlenecks” or stereotypical behaviour be a logical way for systems to evolve to cope with such uncertainty?
  - Need mechanisms which compress prediction uncertainty regularly.
  - Rhythmic interactions might achieve this – chance to realign states ‘on the beat’.
- How do we calibrate such measures of interaction against more subjective notions of interaction?

Using Language

- Language use tends to involve *Joint activities*, composed of *joint actions* requiring *coordination* to reach their mutual goals.
- Evolution of *conventions* to help coordinate.
- Need for *common ground* to have joint activity.
- How can we support this with multimodal interaction?

Basics of Interaction: Joint attention



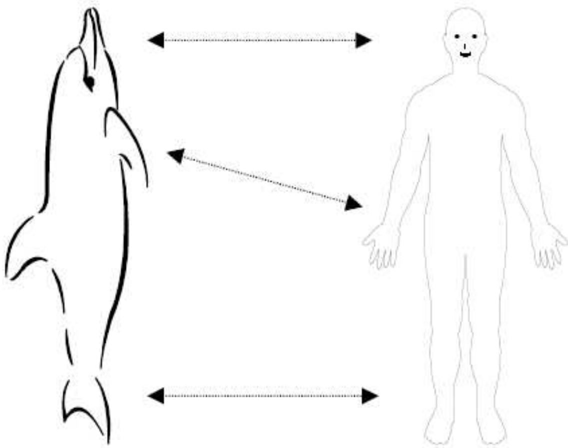
Check attention
Joint engagement

Follow attention
Gaze/ point follow
(Social Referencing)

Direct attention
Imperative & Declarative
Pointing (Referential language)

- We have evolved to participate in collaborative activities involving shared intentionality.
- How much can agent A perceive of the attentional, emotional and motor behaviours of agent B?
- Triadic behaviours involving two people and an object or event about which they share attention
- Viewing other people as intentional agents like themselves.

Basics of Interaction: Imitation



- Imitation is fundamentally linked to language culture and the ability to understand other minds.
  - Becoming a member of a culture means learning new things from other people
- How are actions perceived?
- How do you measure similarity between perception and action?
- As Interaction Designers, how do we help people to solve the *correspondence problem* in remote, multimodal interactions?
  - When users have potentially different devices, with different sensing and display capabilities?

Direct manipulation vs Interface Agents?

- Maybe we should view the interaction more as we do with animals? Think of a rider on a horse, rather than a butler...
  - Rider ‘reads’ the horse, and the horse reads the rider’s intentions via body language, gat, general behaviour, pulling on reins etc
  - Each has its own strengths and weaknesses. Sometimes we need to obey the machine/animal and sometimes it needs to obey us.
  - Smooth Ebb and flow of control between human and machine, rather than a dialogue of instructions and responses.

Outlook

- New challenges for Machine Learning & Inference researchers in HCI
- What is correct balance between display of full posterior distribution of Intention vector, and affordances which suggest control?
- Is it plausible to measure interaction, and can we use that to evolve new systems?

Funding Acknowledgements:

PASCAL
Pattern Analysis, Statistical Modelling and Computational Learning

EPSRC
Engineering and Physical Sciences Research Council

sfi
science foundation ireland
fondúireacht eolaíochta éireann

HEA
Higher Education Authority
An tÚdarás um Ard-Oideachas