Decision Theory: Action Problems

Decision theory goes Bad?

And thus the native hue of resolution
Is sicklied o'er with the pale cast of thought,
And enterprises of great pith and moment
With this regard their currents turn awry,
And lose the name of action.--

Shakespeare-Hamlet

Modeling a Decision task

- Modeling the outcome space
  - What outcomes are likely to influence the decision
- Modeling the event space
  - What events are relevant to the behavior
    - Causal
    - Contextual
- Modeling the beliefs
  - Given relevant event space, determine probabilities
- Modeling utilities
  - Assigning worth to outcomes on a common scale

Perception vs. Action

- Perceptual Utility functions
  - Minimize Errors
  - Possible exception- Geographical slant estimation
- Action Utility functions
  - Minimize energy expenditure (Trajectory selection)
  - Minimize endpoint error (Trajectory selection
  - Maximize information gain (eye/head movements)
  - Minimize collisions (Exploratory navigation)
  - Questions-
    - Facial movements
    - Speech movements
    - Skiing?

Action Decisions

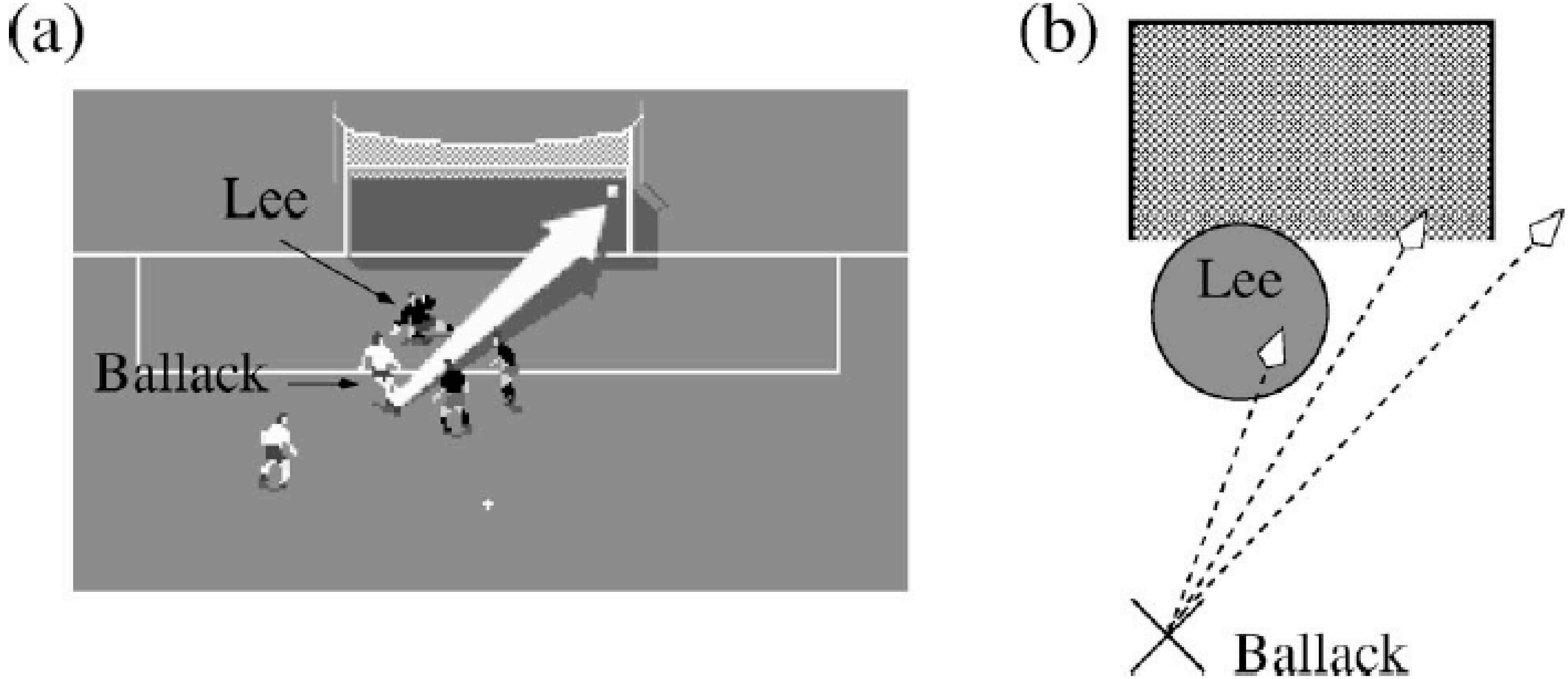


Fig. 1. Michael Ballack's goal during the 2002 World Cup. (a) Ballack, must rapidly decide where to shoot. (b) A schematic of factors affecting the decision.

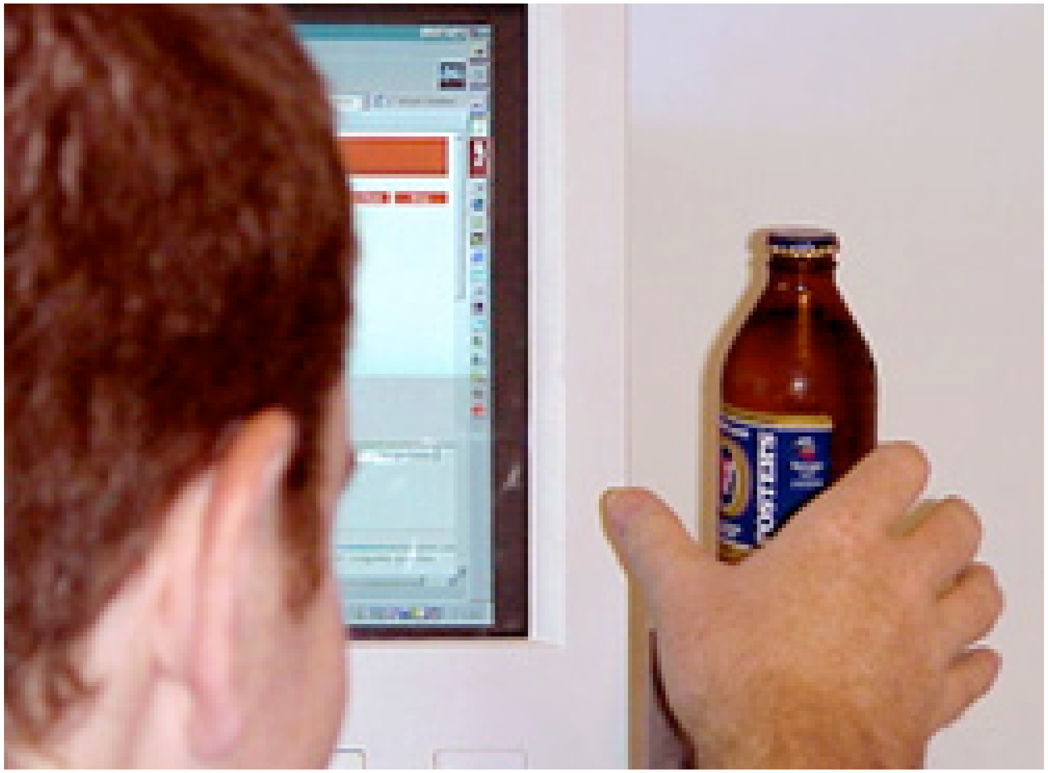
Reaching for an object

Outcomes

State space

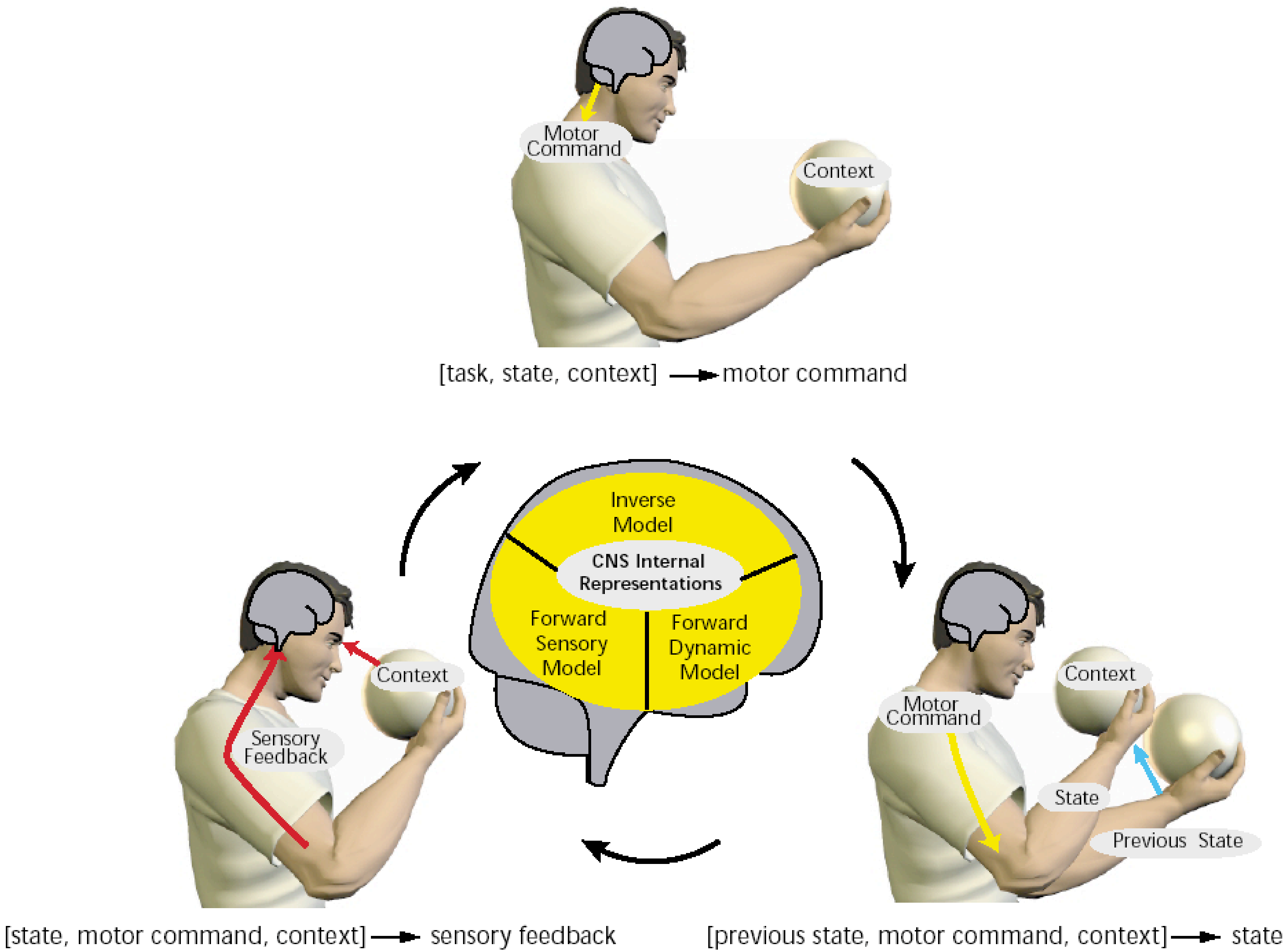
Beliefs

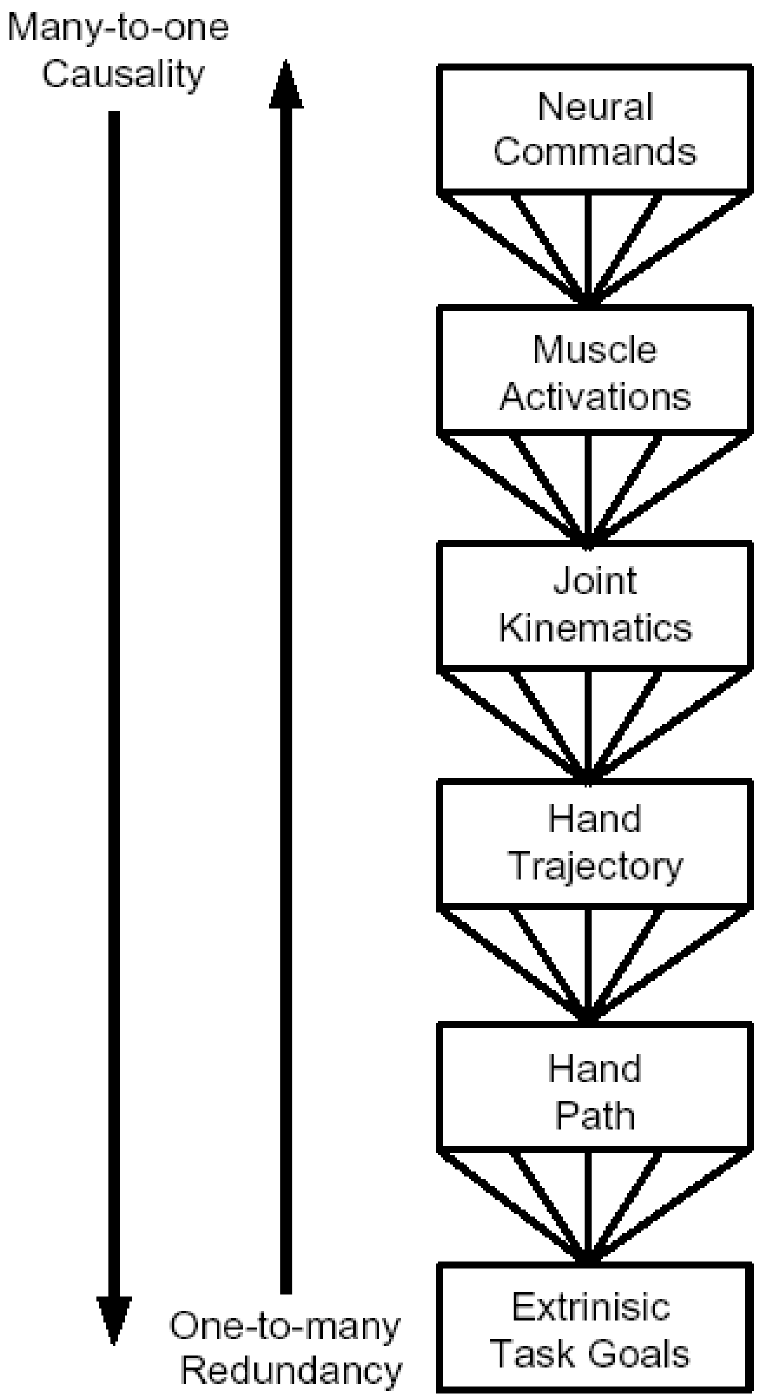
Values



Decision tasks

- Ski downhill
  - Space of outcomes
  - State space
  - Beliefs
  - Values


Motor Command
Context
[task, state, context] → motor command
Inverse Model
CNS Internal Representations
Forward Sensory Model
Forward Dynamic Model
Context
Sensory Feedback
[state, motor command, context] → sensory feedback
Context
Motor Command
State
Previous State
[previous state, motor command, context] → state



Motor Control is Hierarchical

The levels in the motor hierarchy are shown with the triangles between the levels indicating the reduction in the degrees of freedom between the higher and lower levels. Specifying a pattern of behavior at any level completely specifies the patterns at the level below (many-to-one: many patterns at the higher level correspond to one pattern at the lower) but is consistent with many patterns at the level above (one-to-many). Planning can be considered as the process by which particular patterns, consistent with the extrinsic task goals, are selected at each level. From (Wolpert, 1997).

Movements show typicality

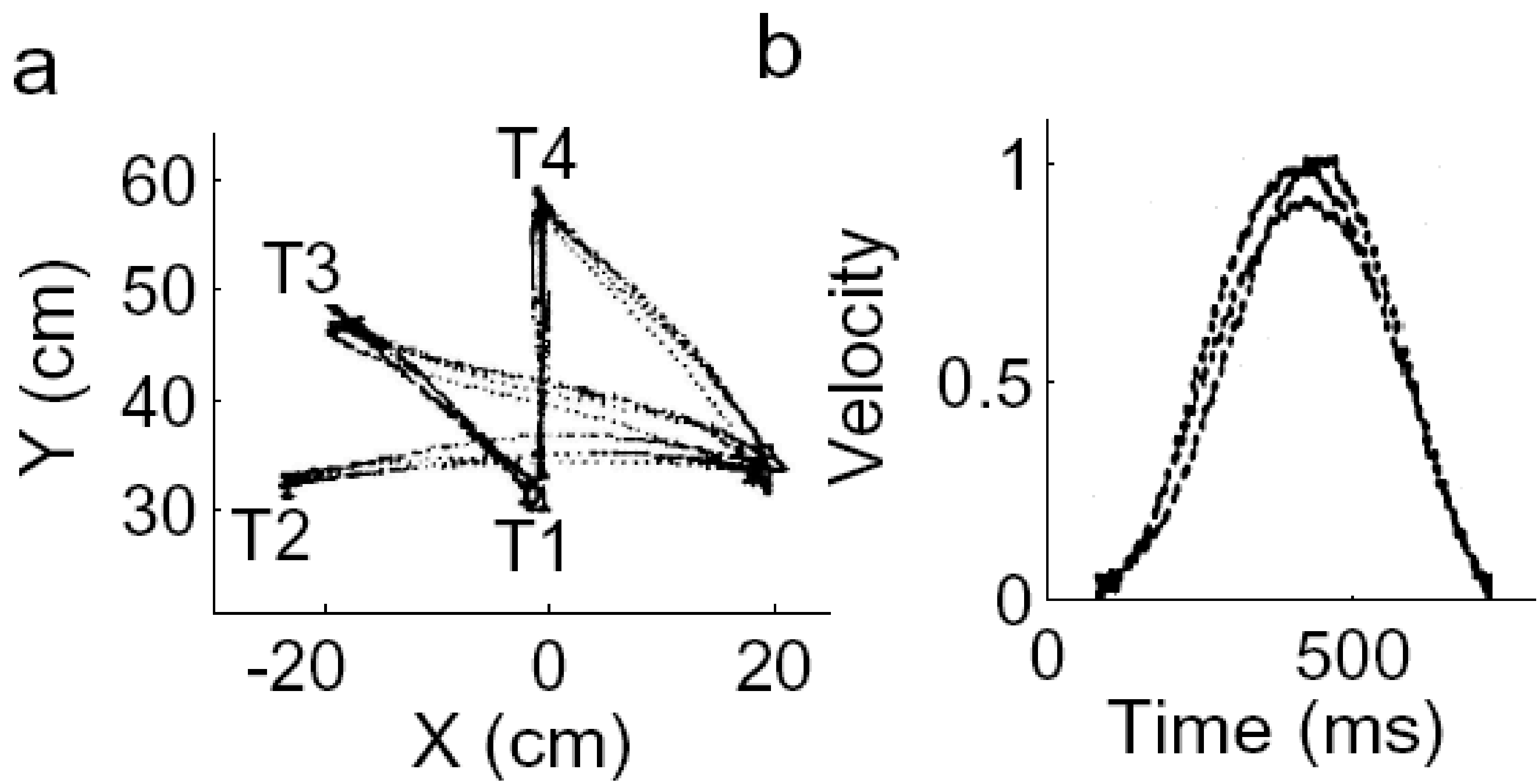
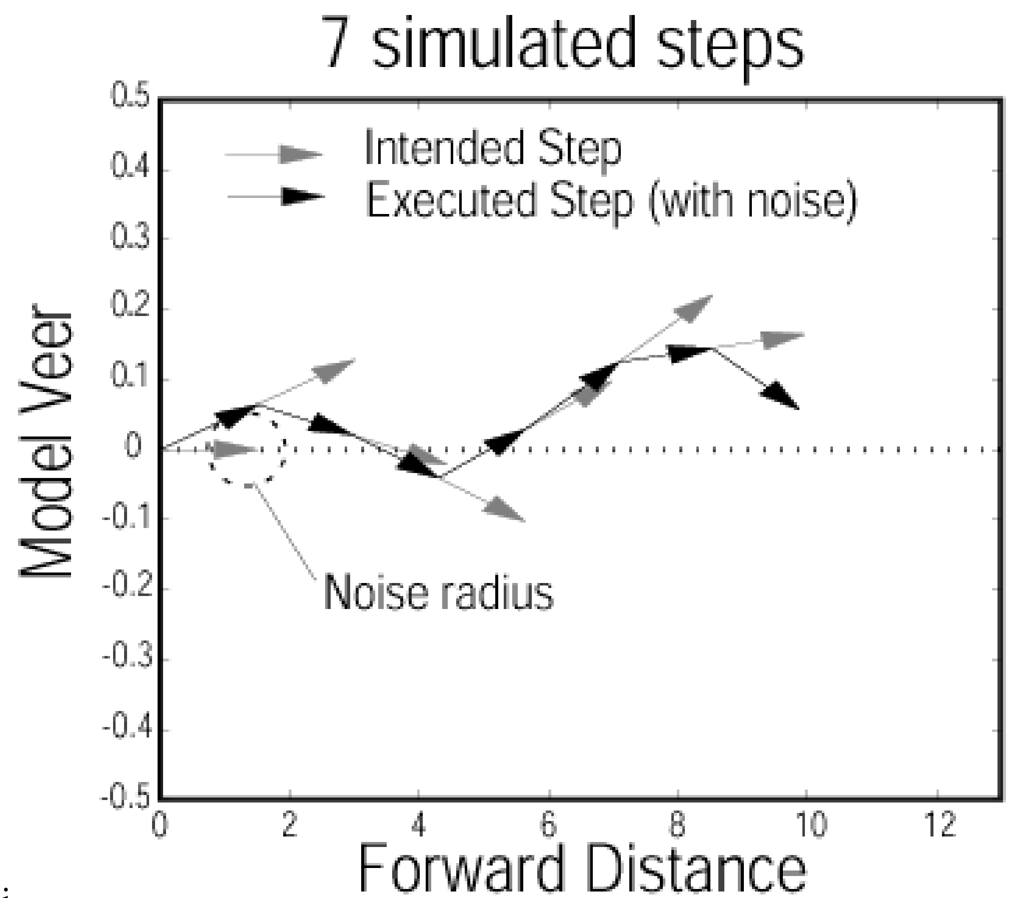


Figure 3. a) observed hand paths for a set of point-to-point movements from (Uno et al., 1989) (with permission). The coordinate system is centred on the shoulder with x and y in the transverse and sagittal directions respectively. b) observed velocity profiles for movements from T1 to T3 in a). Reprinted with permission from (Uno et al., 1989)

Optimal control theory as decision theory with dynamics (sequential decision problem)

Decisions may occur at each time step in a movement. Thus the utility function must be specified at each time.



Utility of action sequences

Best movements maximize total utility

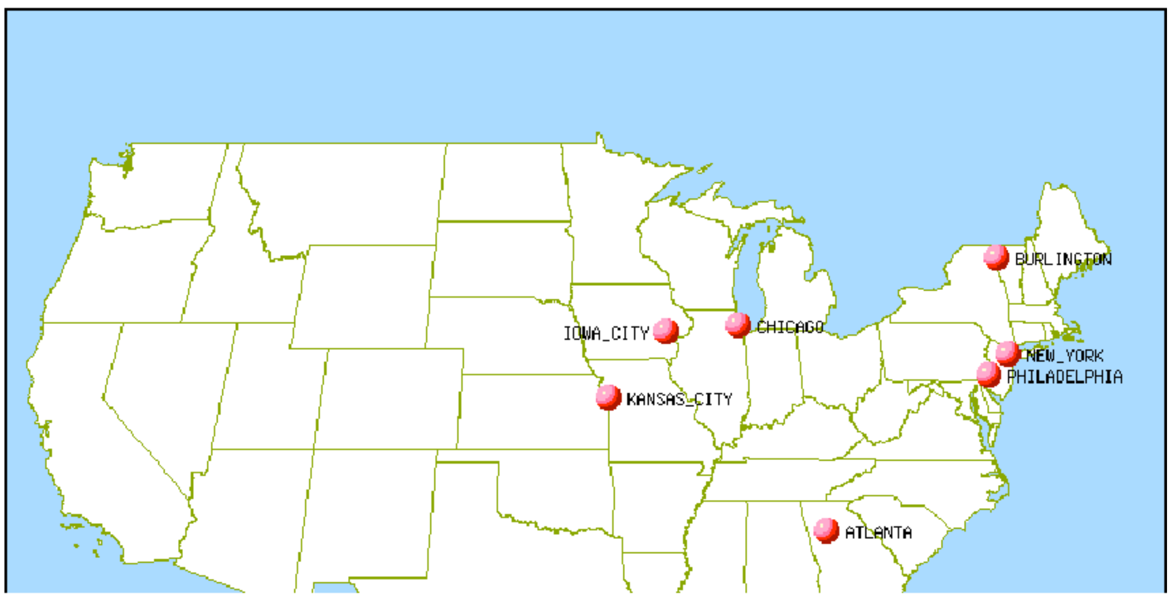
– every possible movement which can achieve a goal has a utility

– we select the movement with the highest utility

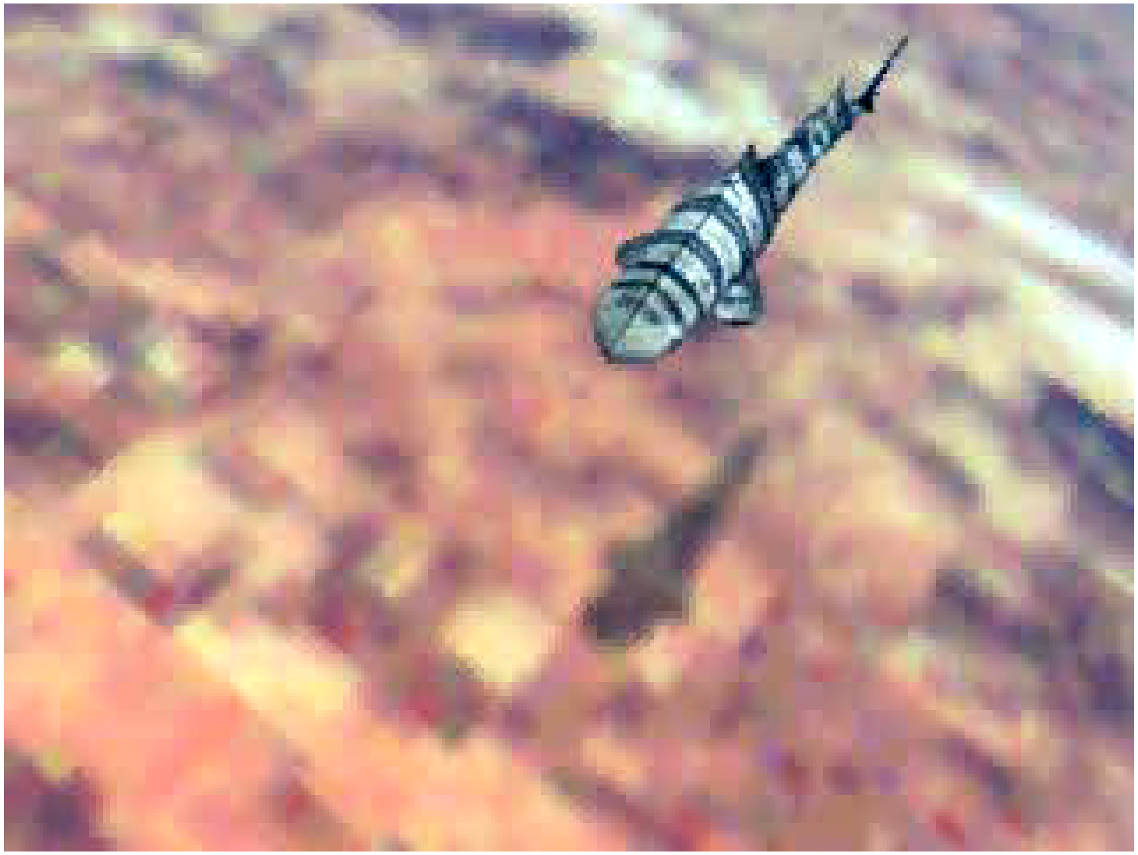
Traveling Salesman problem example:

	1	2	3	4	5	6	7	8	9
1 Chicago IL		203	748	954	579	712	672	997	400
2 Iowa City IA	203		946	858	665	914	873	1070	222
3 Burlington VT	748	946		1595	938	264	317	1240	1147
4 Houston TX	954	858	1595		727	1433	1368	807	670
5 Atlanta GA	579	665	938	727		732	666	418	676
6 New York NY	712	914	264	1433	732		66	992	1085
7 Philadelphia PA	672	873	317	1368	666	66		929	1036
8 Tampa FL	997	1070	1240	807	418	992	929		1039
9 Kansas City MO	400	222	1147	670	676	1085	1036	1039	

Cities Entered
Chicago
Iowa City
Burlington
Houston
Atlanta
New York
Philadelphia
Tampa
Kansas City



Minimal Energy Models for movement



Previous (incorrect) costs

Saccadic eye movements

- little vision over 4°/s
- saccades >200°/s
- frequent 2-3 /sec each ~50ms
- deprives of vision ~90 min/day

⇒Minimize time

Arm Movements

- Are smooth

⇒ Minimum jerk
(rate of change of acceleration)

⇒ Minimum torque change

Criteria for cost for goal-directed movement

- Makes sense in terms of advantage for evolution & learning
- Simple for CNS to measure
- Generalizes to different systems e.g. eye, head, arm
- Generalizes to different tasks e.g. pointing, grasping, drawing

→ Reproduce & predict behaviour

Utility Functions for reaching

Simple utility function- Minimum *Jerk*

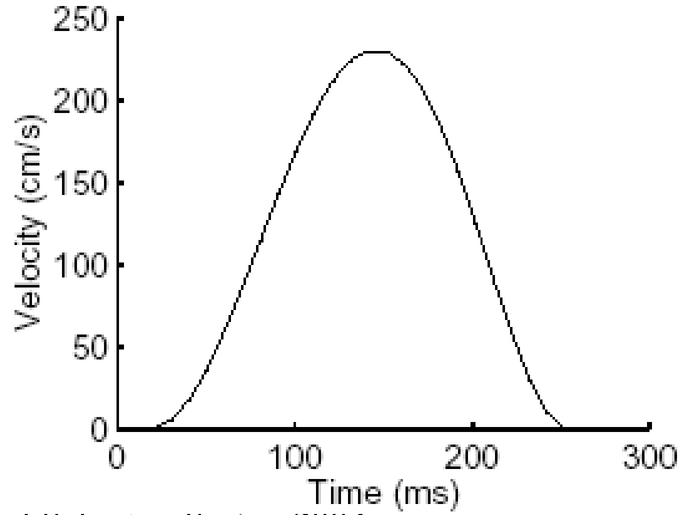
Solution has the form:

$$J = \int_0^T \left[\frac{d^3x}{dt^3}\right]^2 dt$$

$$x(t) = x_0 + (x_f - x_0)[10(t/T)^3 - 15(t/T)^4 + 6(t/T)^5]$$

Model predicts bell-shaped velocity profile.

No role for uncertainty.



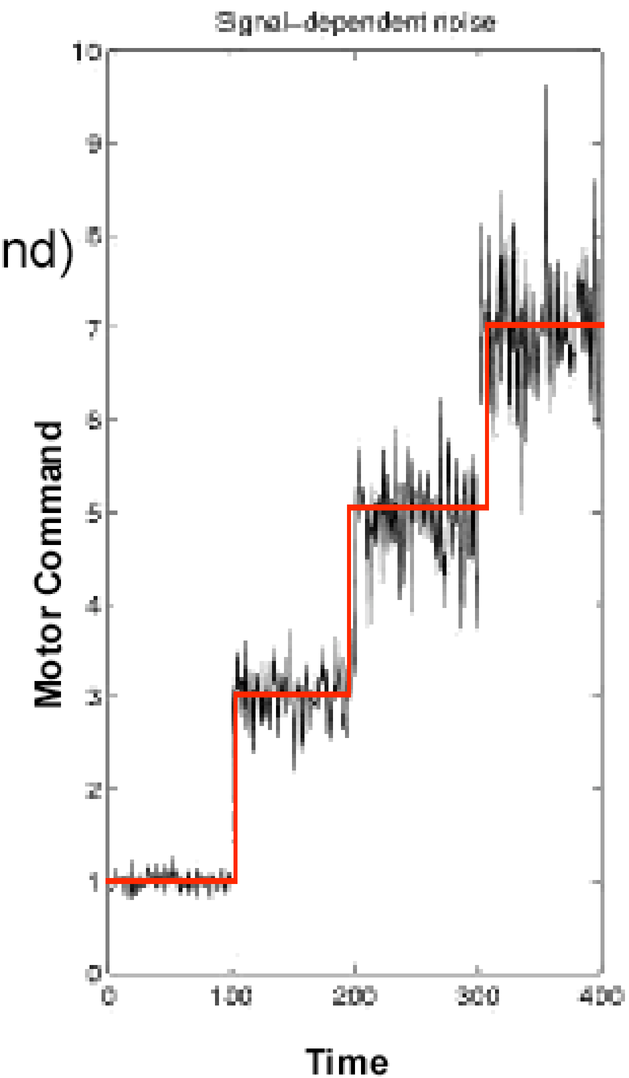
Task/Goal Achievement?

Fundamental constraint=Signal-dependent noise

- Signal-dependent noise:
  - Constant coefficient of variation
  - SD (motor command) ~ Mean (motor command)

- Evidence from
  - Experiments: SD (Force) ~ Mean (Force)
  - Modelling
    - Spikes drawn from a renewal process
    - Recruitment properties of motor units

(Jones, Hamilton & Wolpert , J. Neurophysiol., 2002)



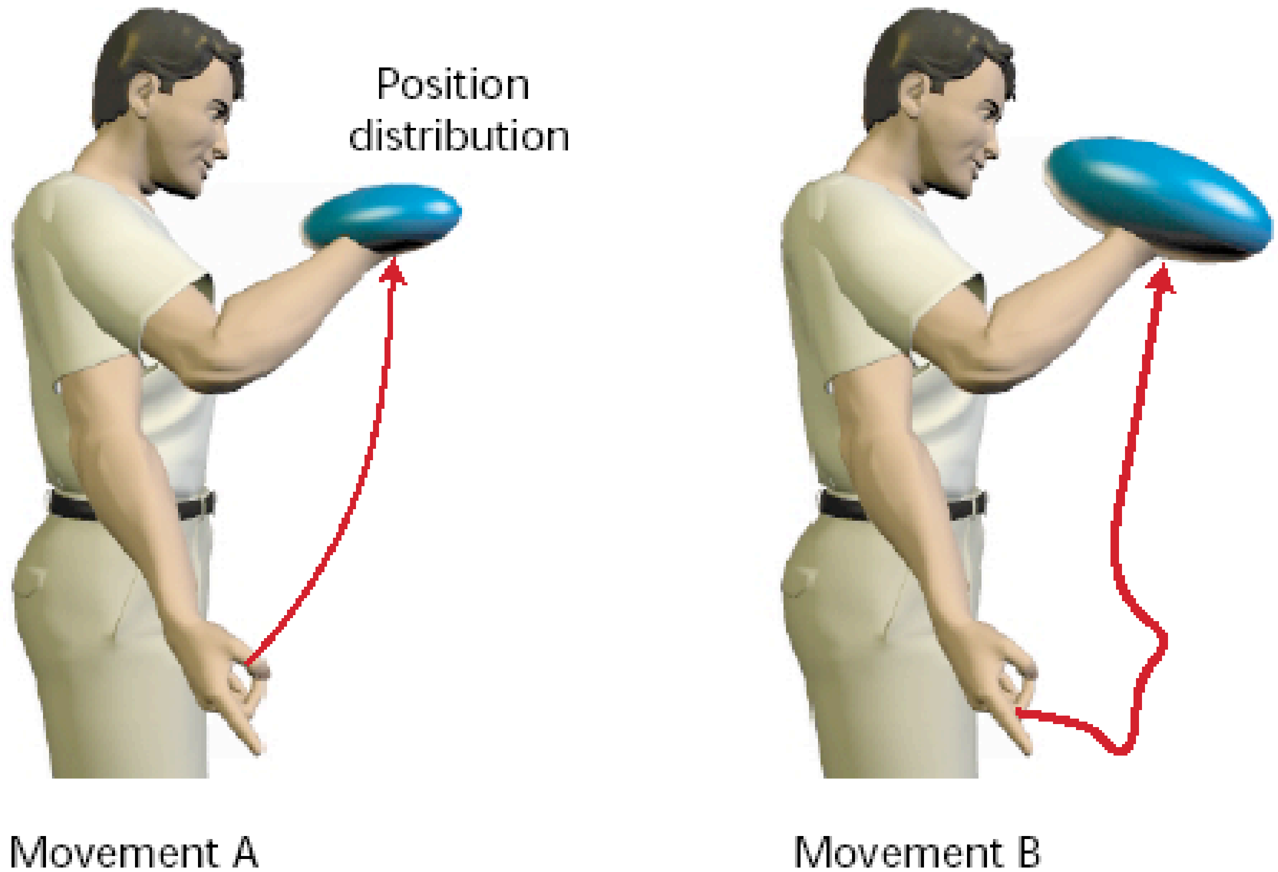
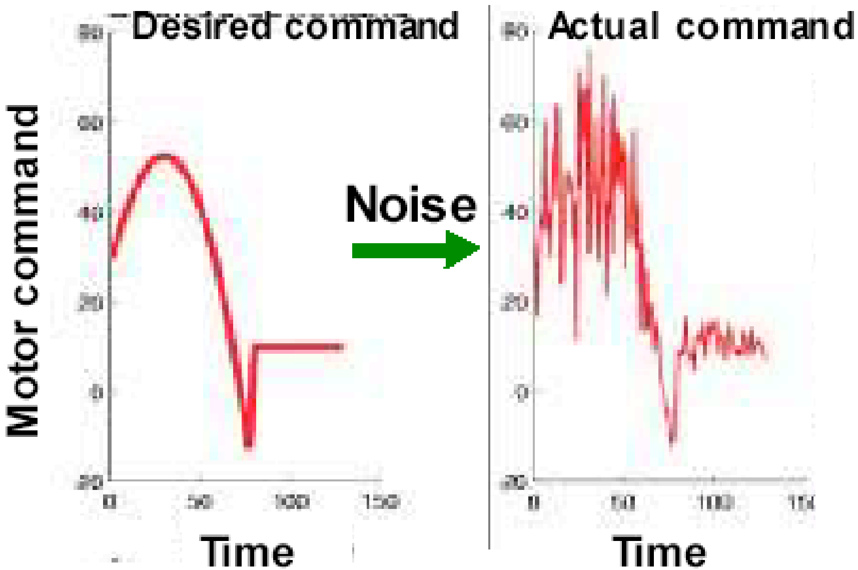
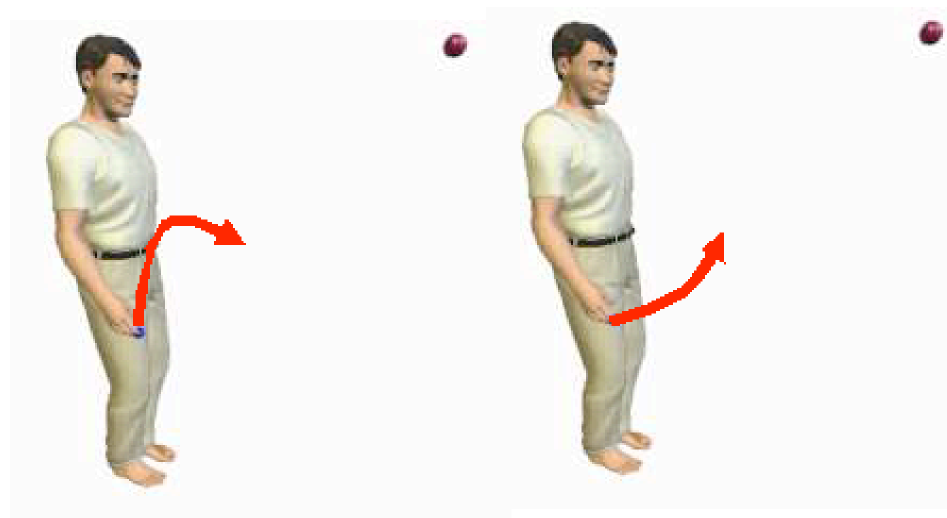


Fig. 2. Task optimization in the presence of signal-dependent noise (TOPS) model of Harris and Wolpert[9]. Average paths and expected final position distributions for two different motor sequences. Although the sequences bring the hand on average to the same final position, they have different final distributions because of noise in the motor commands. Movement A has smaller spread than B and therefore has lower cost than B. In general, the task determines the desired statistics of the movement, and the trajectory that optimizes the statistics is selected.

PSY

Signal-dependent noise and task achievement





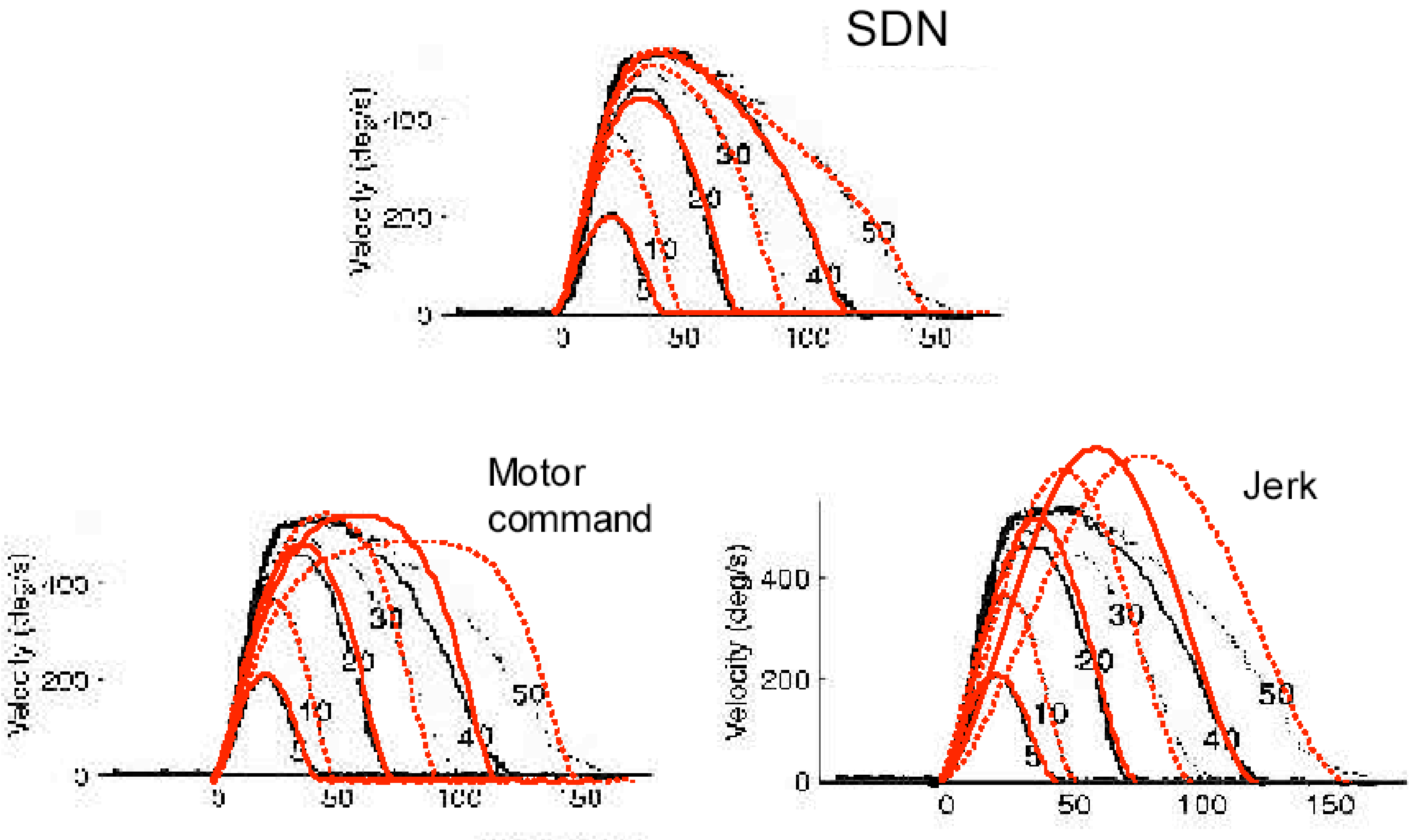
Sequence of motor commands $\Rightarrow$ probability distribution (statistics) of movement.

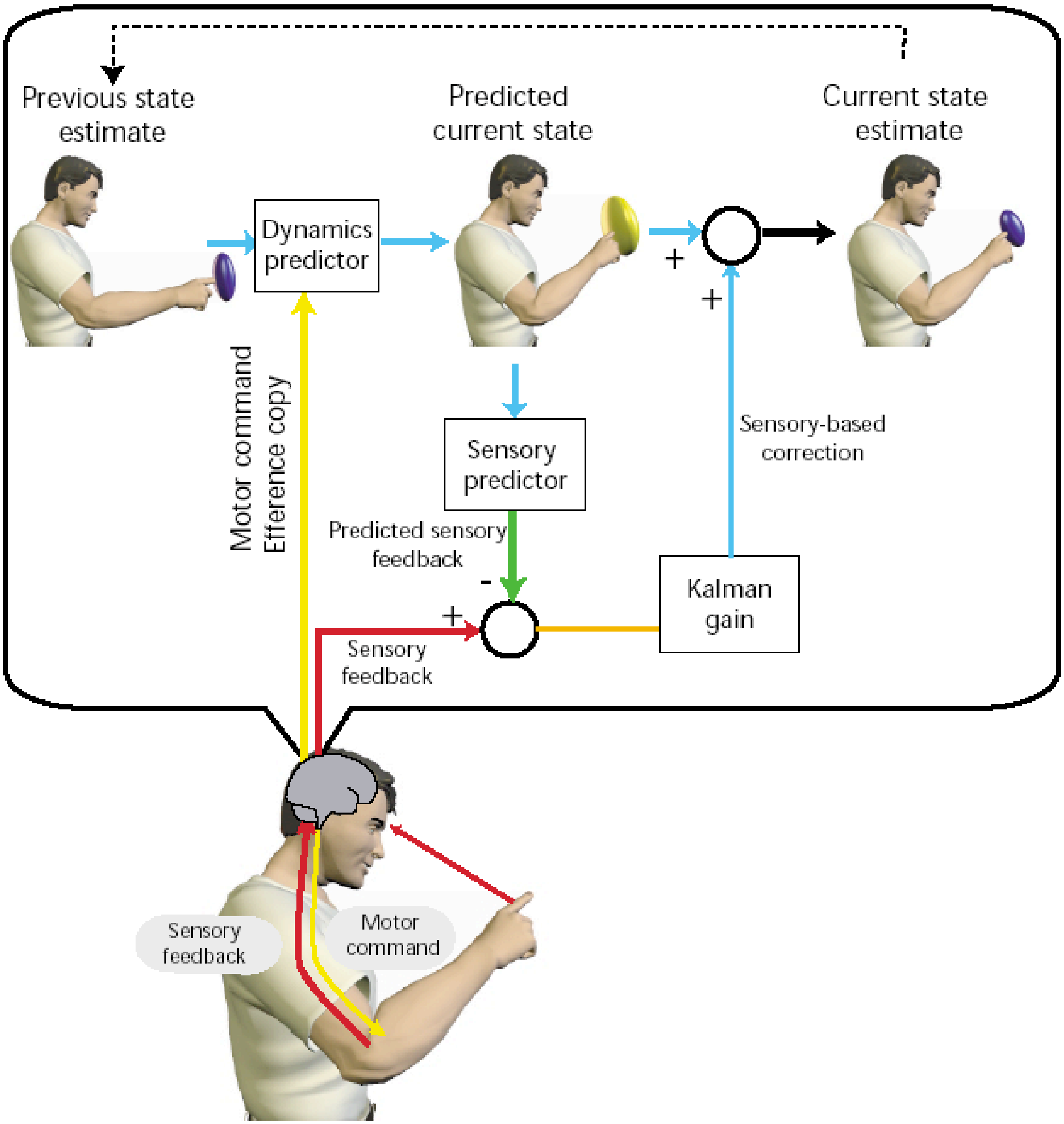
The statistics of action can be controlled by changing the motor command

Task $\equiv$ Optimizing f(statistics)

(Harris & Wolpert, Nature, 1998)

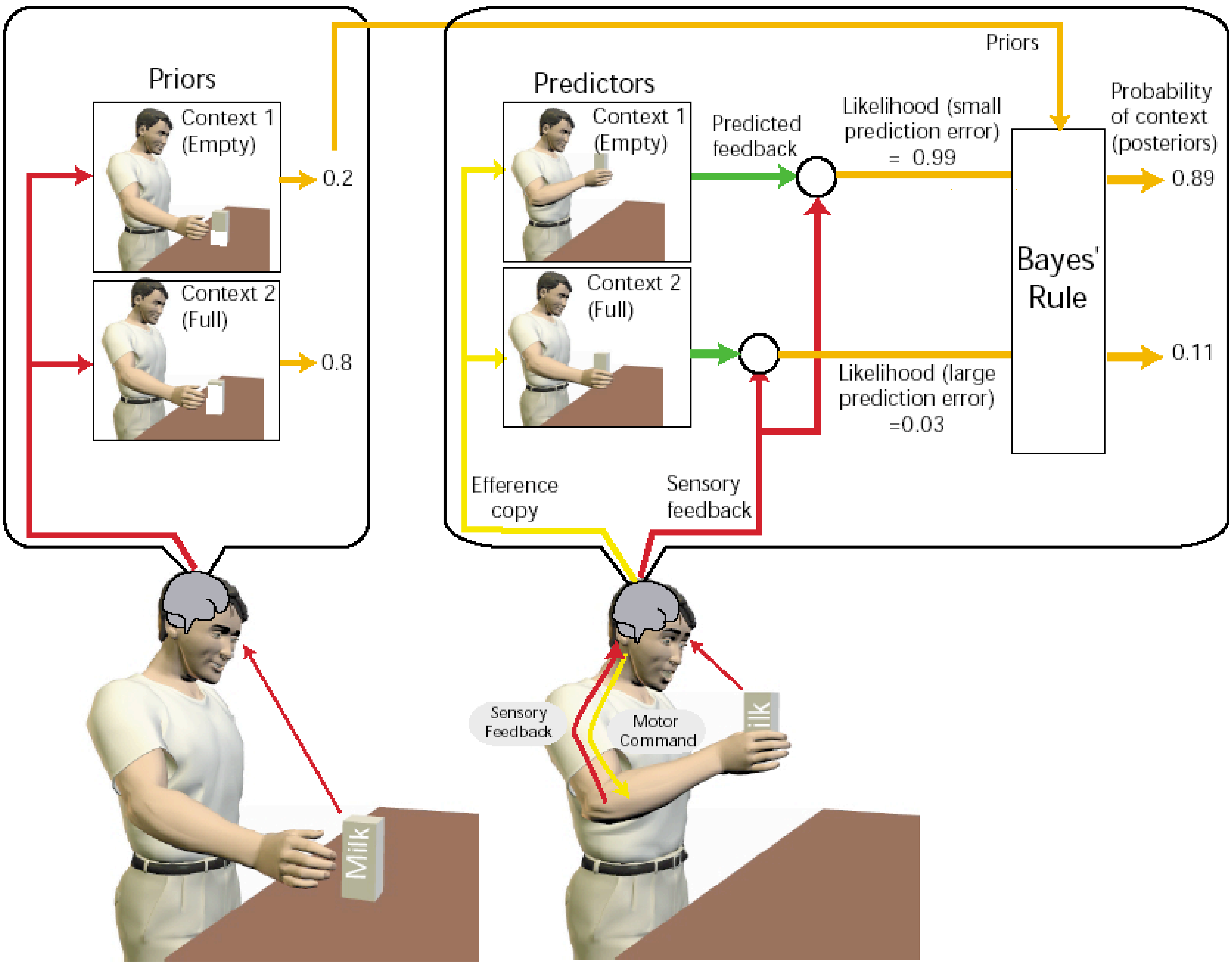
Saccade predictions



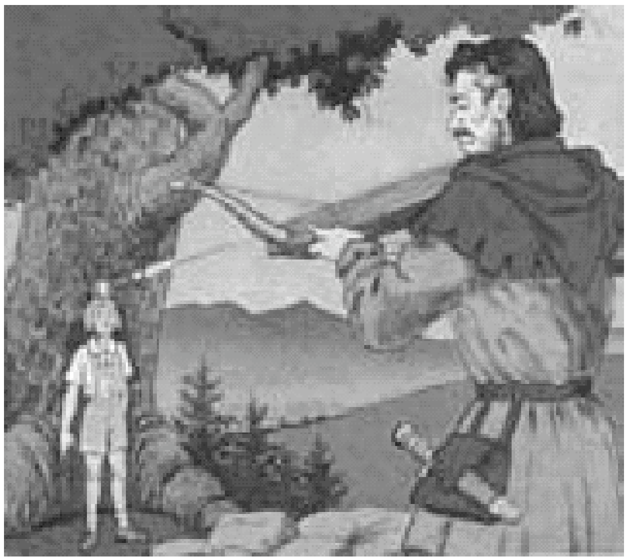


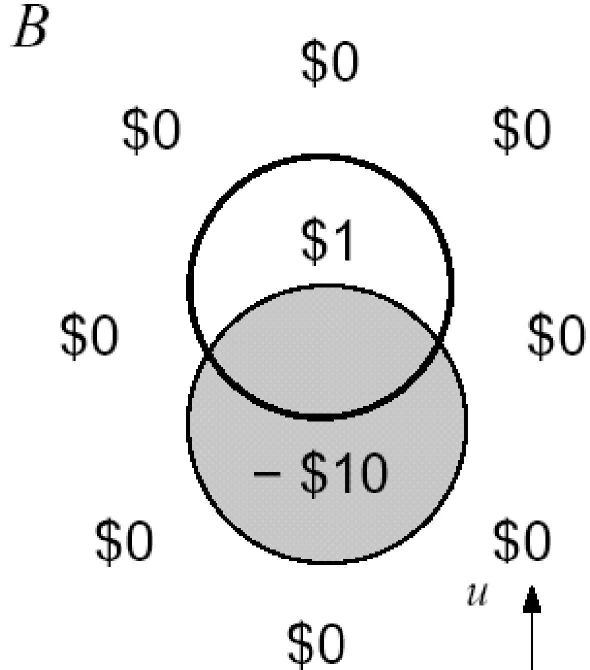
Previous state estimate
Dynamics predictor
Predicted current state
Current state estimate
+
+
Motor command
Efference copy
Sensory predictor
Sensory-based correction
Predicted sensory feedback
-
+
Kalman gain
Sensory feedback
Sensory feedback
Motor command

PS


Priors
Context 1
(Empty)
0.2
Context 2
(Full)
0.8
Priors
Predictors
Context 1
(Empty)
Context 2
(Full)
Predicted
feedback
Likelihood (small
prediction error)
= 0.99
Likelihood (large
prediction error)
=0.03
Bayes'
Rule
Probability
of context
(posteriors)
0.89
0.11
Efference
copy
Sensory
feedback
Milk
Sensory
Feedback
Motor
Command

Projectile Actions





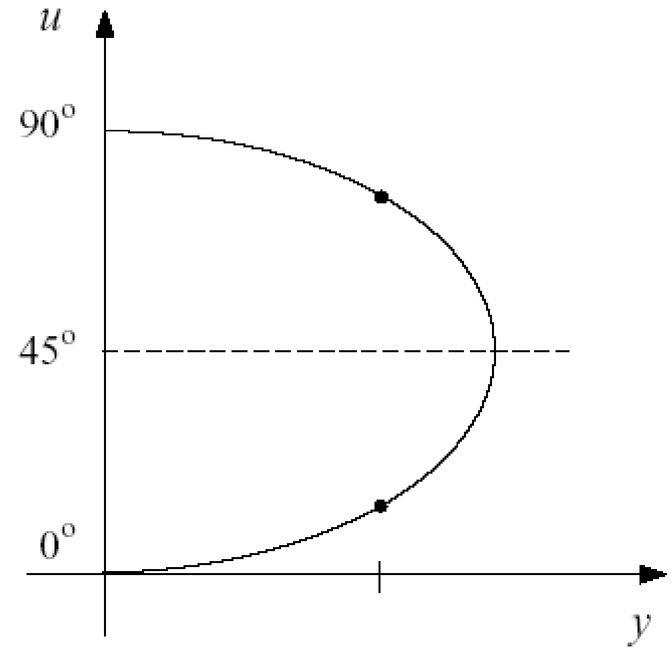


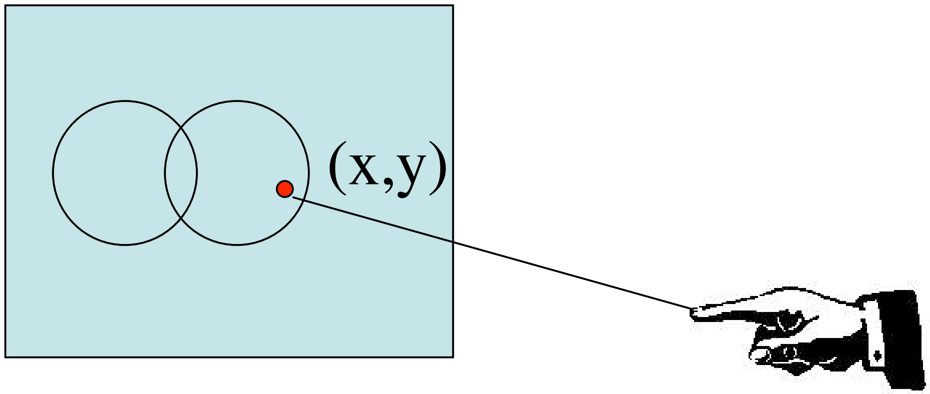
Figure 8. (a) An "archery" problem. (b) The parabolic relationship between distance traveled (y) and angle (u) for a projectile. For each value of y there are two corresponding values of u, symmetrically placed around 45 degrees.

Model State space

- Reach endpoint
- Reach trajectory?

- Model beliefs on endpoint
  - Planned endpoint plus 2-D gaussian noise

Model Outcomes



Action = (x,y)

Land in circle 0: R0
Land in circle 1: R1
Reach too long: timeout
Energy for reach: B(x,y)

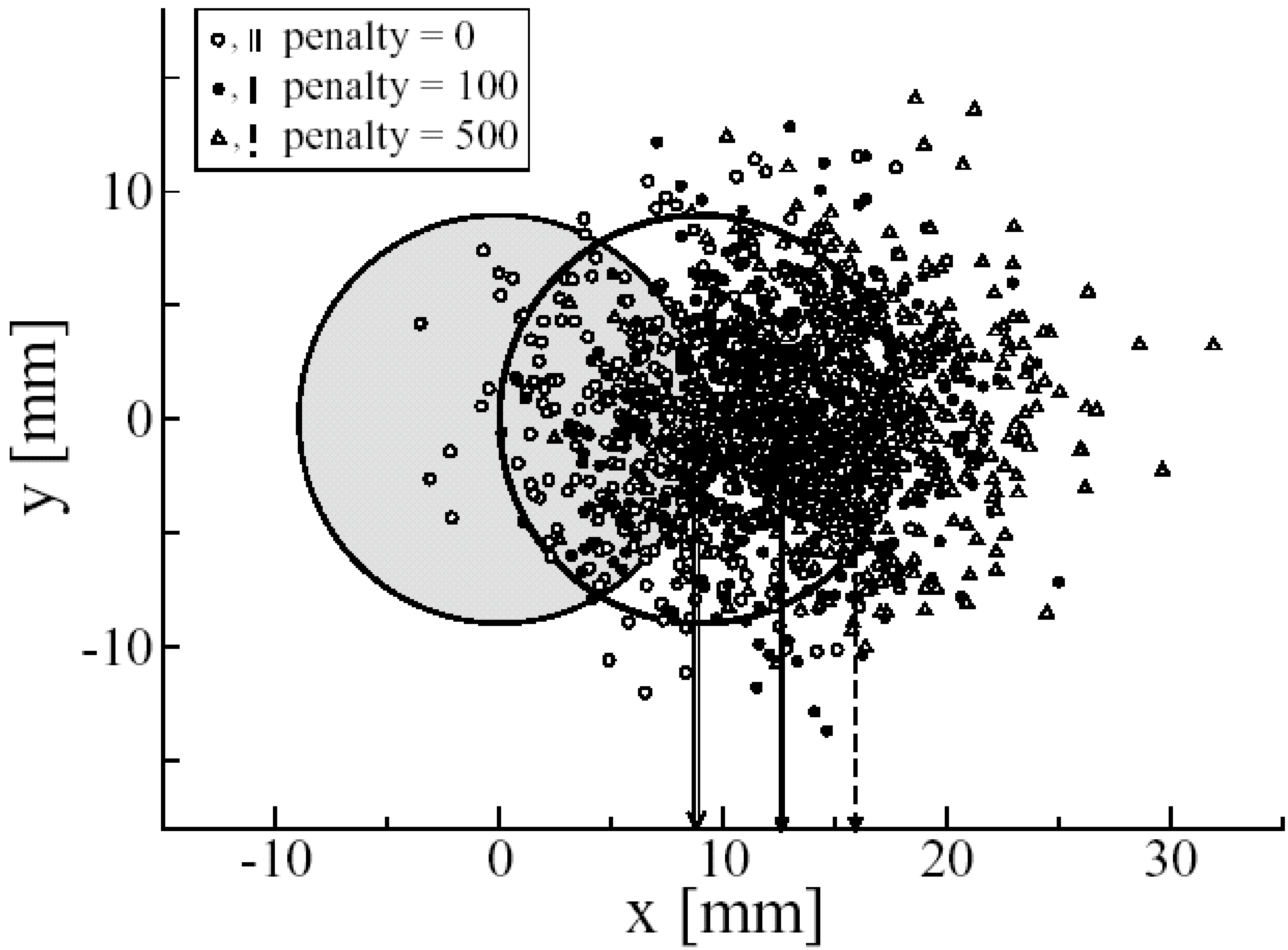
Multi-Attribute Utility:

- Energy
- Timeout
- Rewards

Minimize Expected Utility:

$$
\begin{aligned}
V(x,y) &= U(Ro)P(Ro \mid x,y) + U(R_1)P(R_1 \mid x,y) \\
&\quad + U(timeout)P(timeout \mid x,y) + U(B(x,y))P(B \mid x,y)
\end{aligned}
$$

Simulate Optimal Pointer



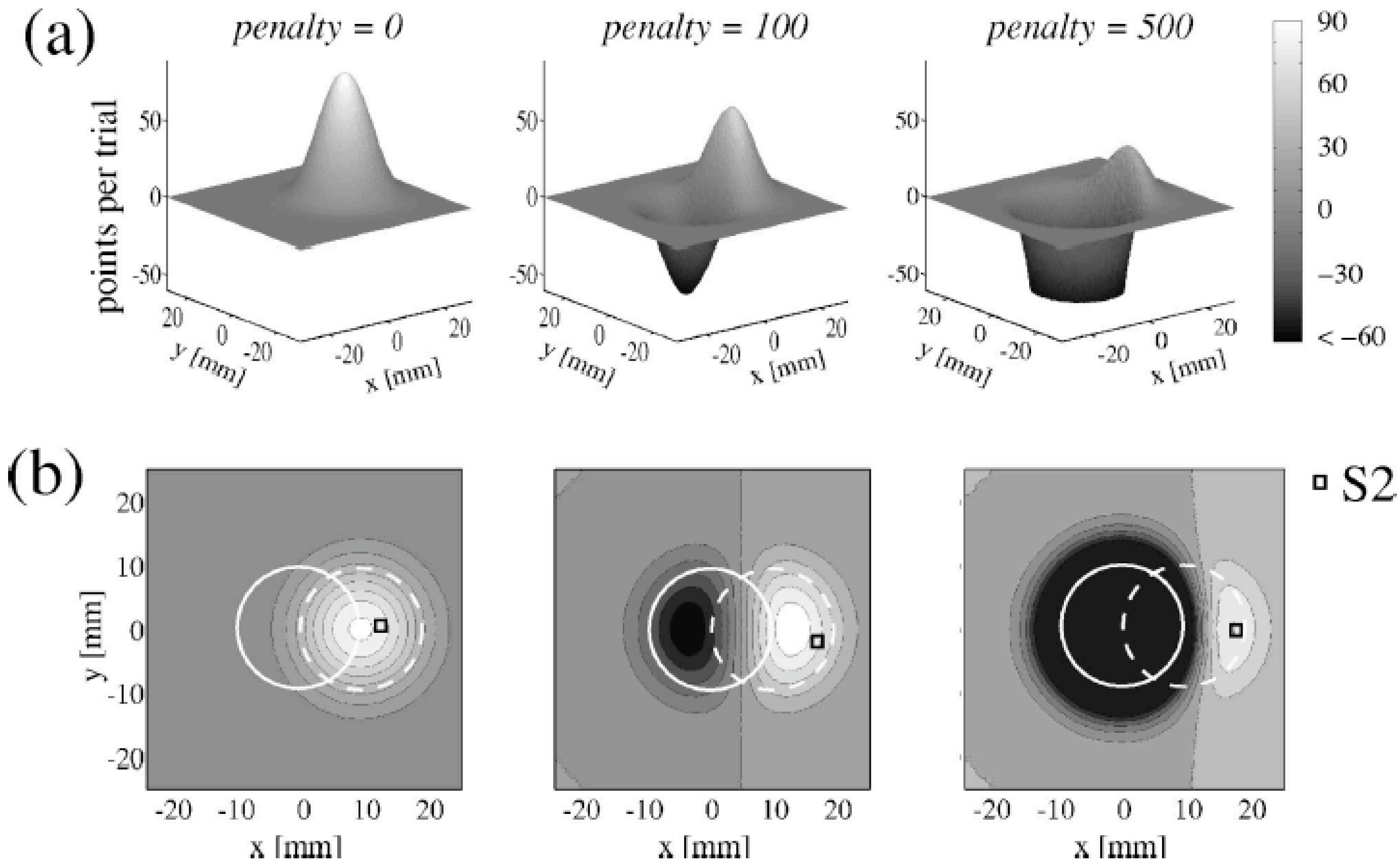
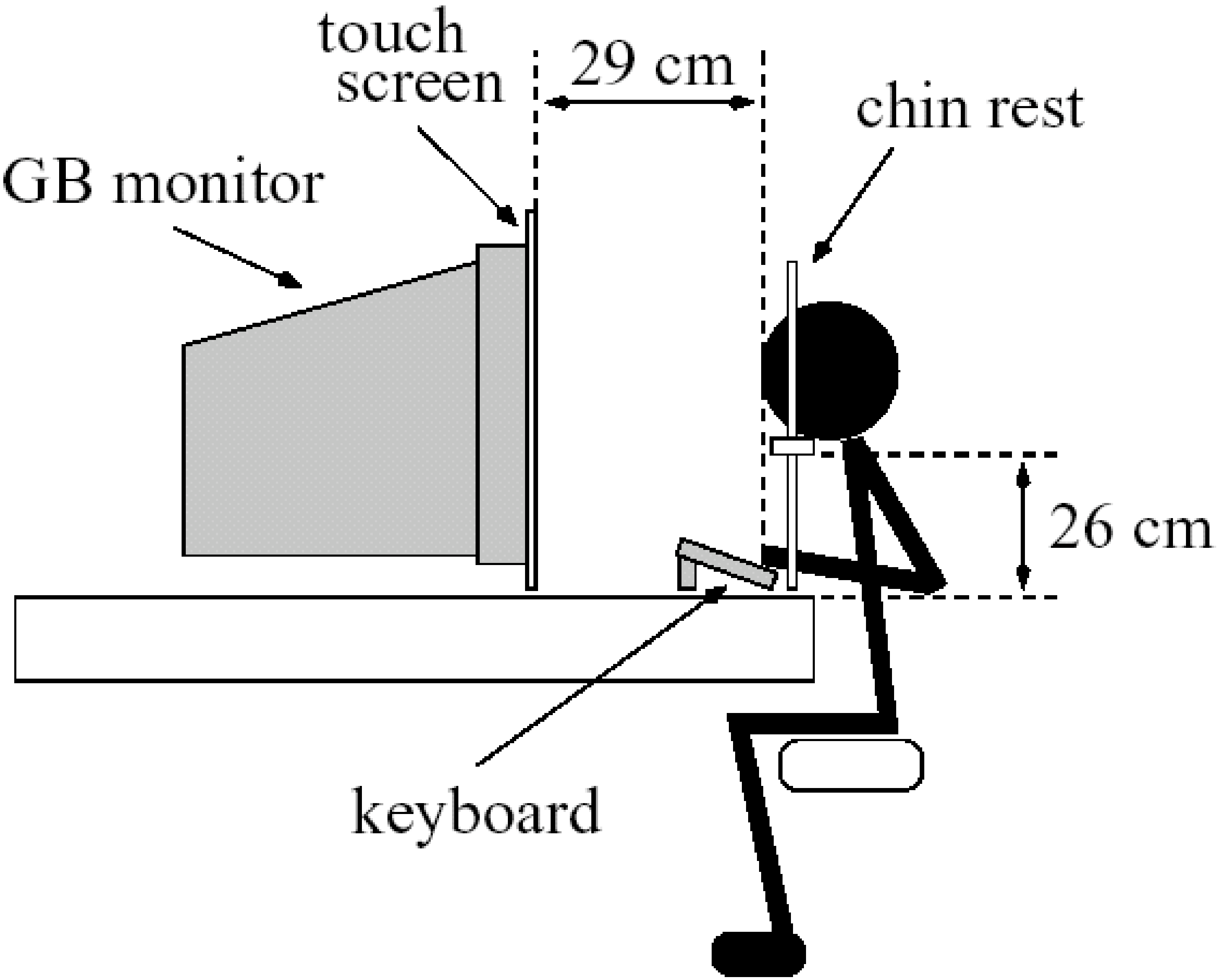
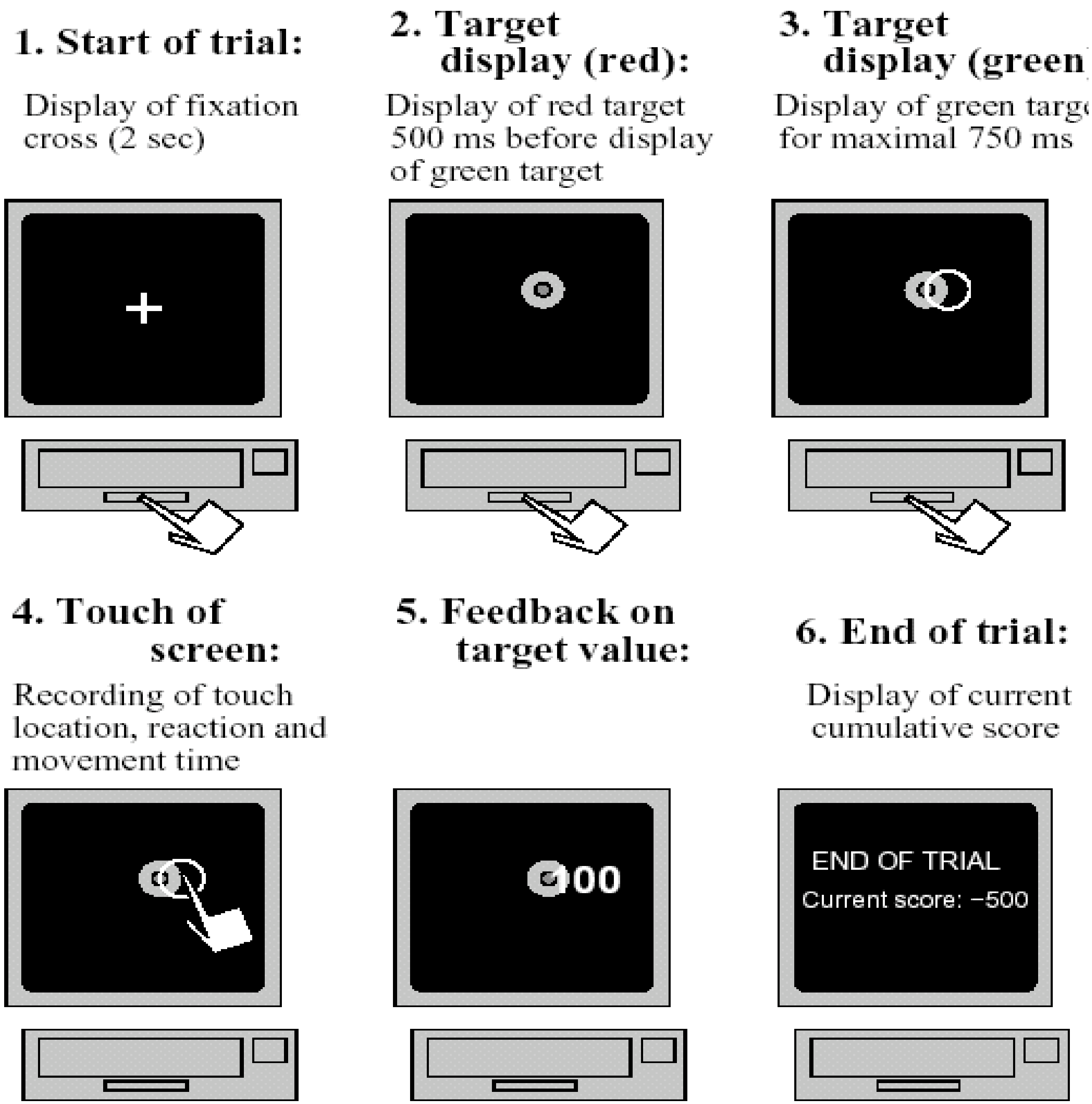


Fig. 2. "Landscape" of expected gain for an optimal observer with a variance of $\sigma^2 = 4.83$ (matching that of subject S2 in experiment 1). (a) Expected gain (in points per trial) as a function of the mean movement end point (x, y). The distribution is truncated for scores < -60 points. (b) The same landscape replotted as a contour plot with the mean movement end point of subject S2 (open squares) compared with optimal performance as predicted by the model [the contour regions are coded with the same gray-level scale as in (a)].

Task - Touch the screen




1. Start of trial:
Display of fixation cross (2 sec)
2. Target display (red):
Display of red target 500 ms before display of green target
3. Target display (green
Display of green targ
for maximal 750 ms
4. Touch of screen:
Recording of touch location, reaction and movement time
5. Feedback on target value:
00
6. End of trial:
Display of current cumulative score
END OF TRIAL
Current score: −500

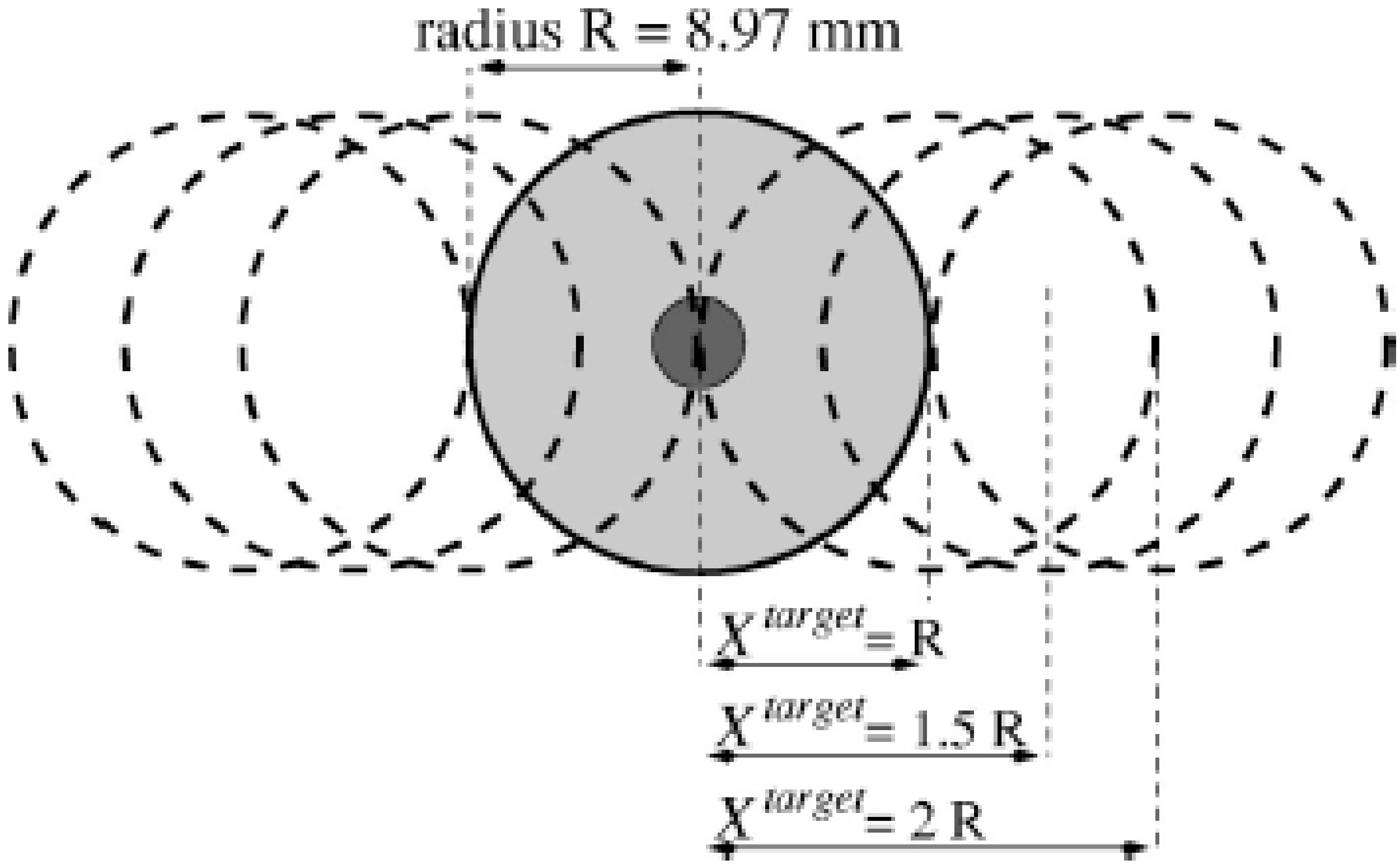
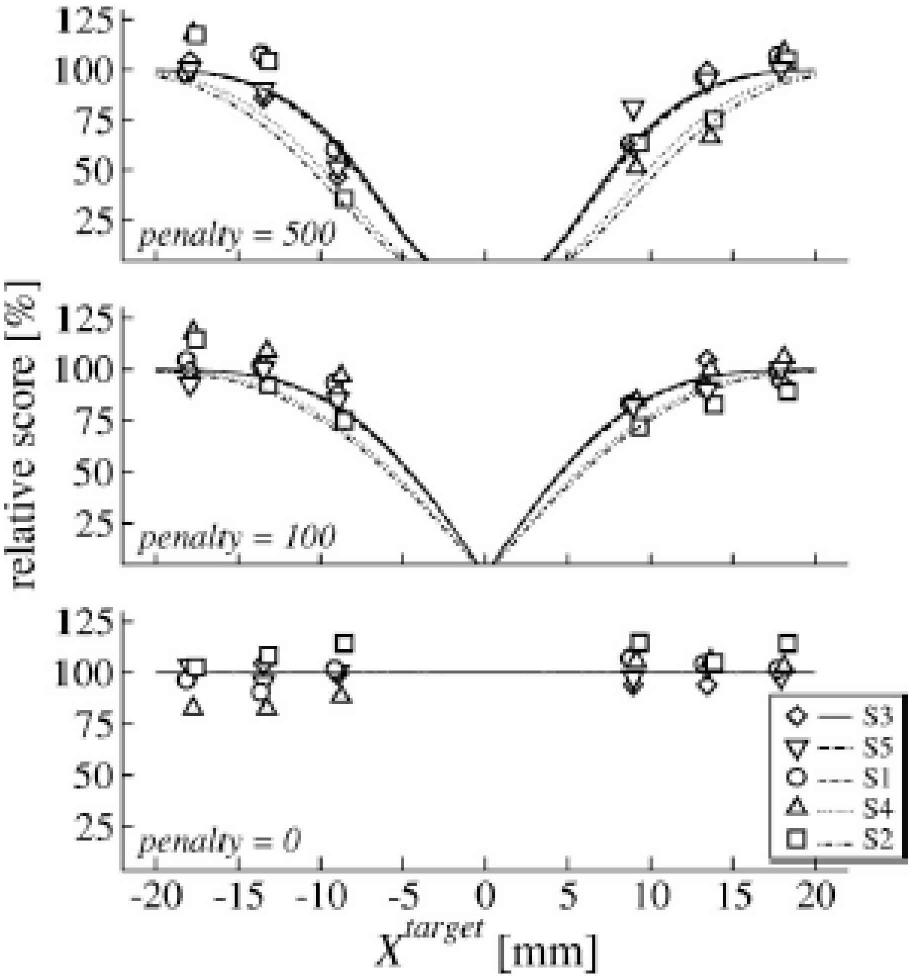


Fig. 3. Layout of the stimuli in experiment 1. The six dashed regions indicate the six different positions at which the target could appear.



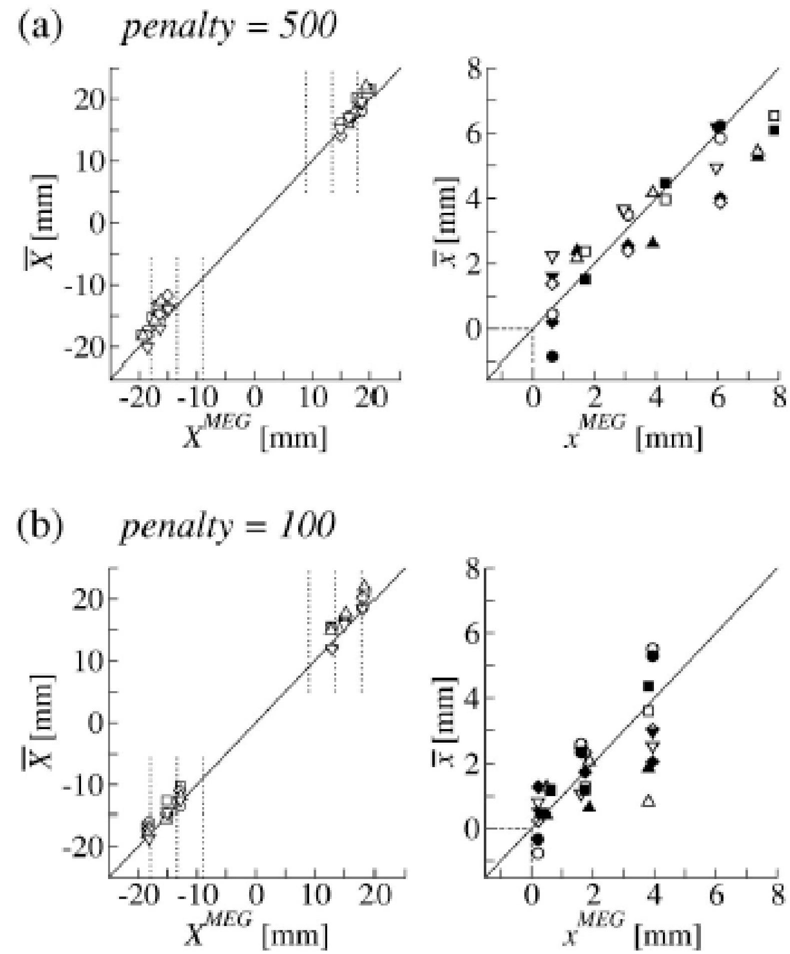


Fig. 6. Experiment 1, results for five subjects, listed in order of motor variability, for penalty conditions 0, 100, and 500. The values plotted on the vertical axis are average scores per target position displayed as a percentage of optimal performance predicted by the model for penalty = 0. Normalizing in this way makes it easier to compare performance of subjects with different motor variabilities. The horizontal axis is the target position X^{target} relative to the penalty region. Model predictions based on each subject's variability were computed. The curves (one per subject) represent the model predictions.

Eye movements

- Outcome space
  - End point accuracy- foveate target
  - Acquire relevant target information
- Event space
  - Eye position
  - Target identity

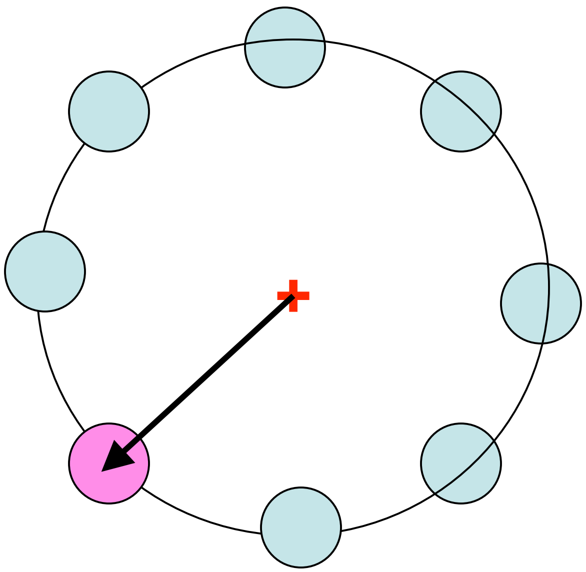
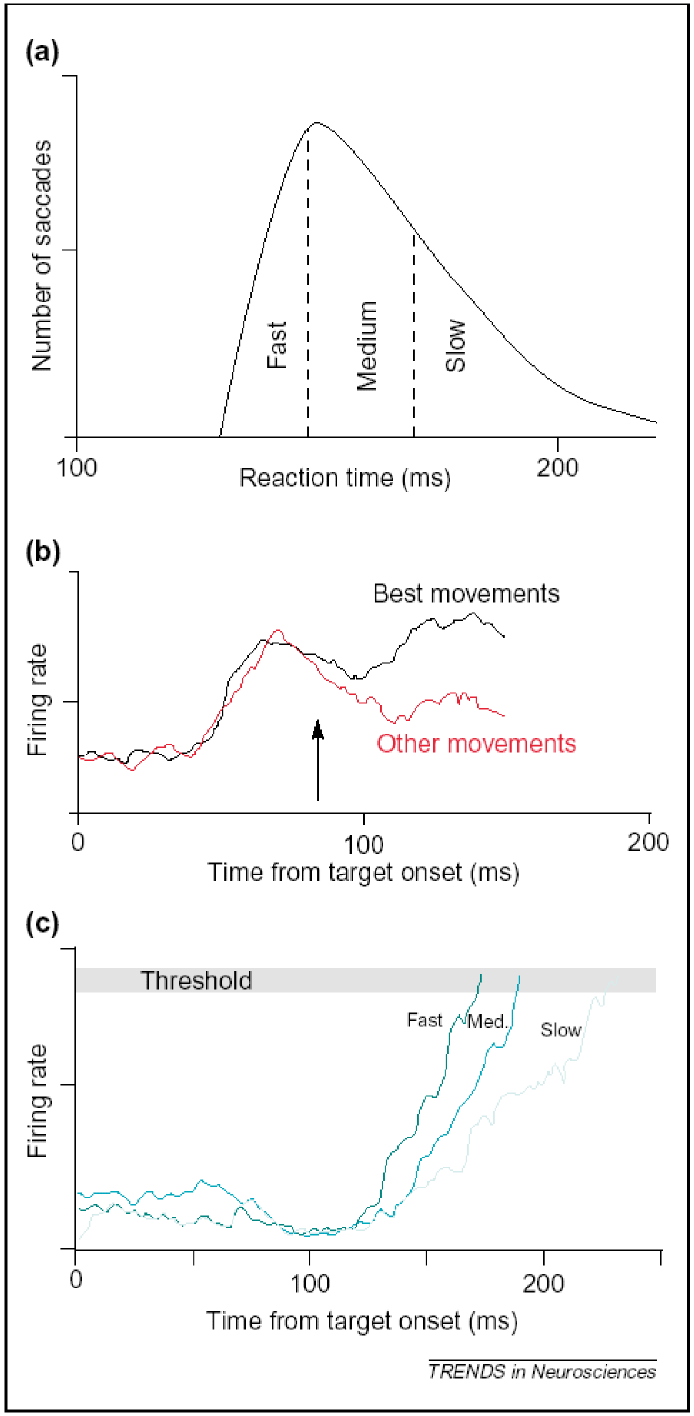
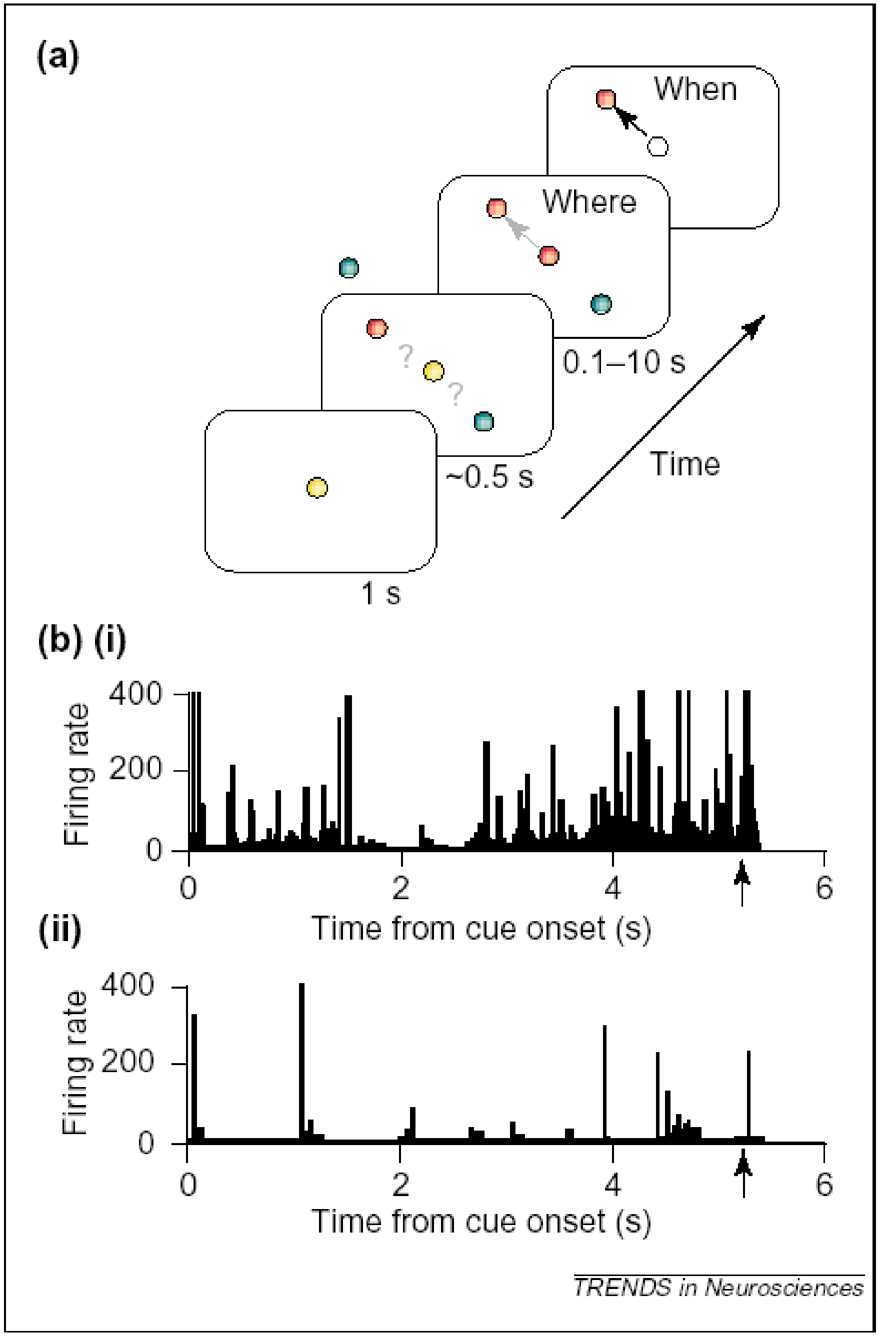


Fig. 2. A profile of decision-related activity in the frontal eye fields during a simple oddball detection task. (a) Monkeys detect the oddball and shift gaze towards it with a variety of reaction times. (b) The black line plots average firing rate for frontal eye field neurons during trials on which the location of the oddball target was positioned to elicit the best movement of the neuron under study. The red line plots average firing rate on trials in which the oddball was positioned to elicit movements for which the neuron was unmodulated. For movements with fast reaction times, neuronal firing rates on these two types of trials begin to diverge about 80 ms after target onset. Surprisingly, for medium and slow movements, firing rates for best movement and other movement trials also diverge at about 80 ms (data not shown). (c) When average firing rates on best movement trials are plotted for fast, medium and slow movements, it is found that movements (data not shown) begin a fixed interval after the neurons reach a crucial firing rate or threshold. Adapted, with permission, from Ref. 10.





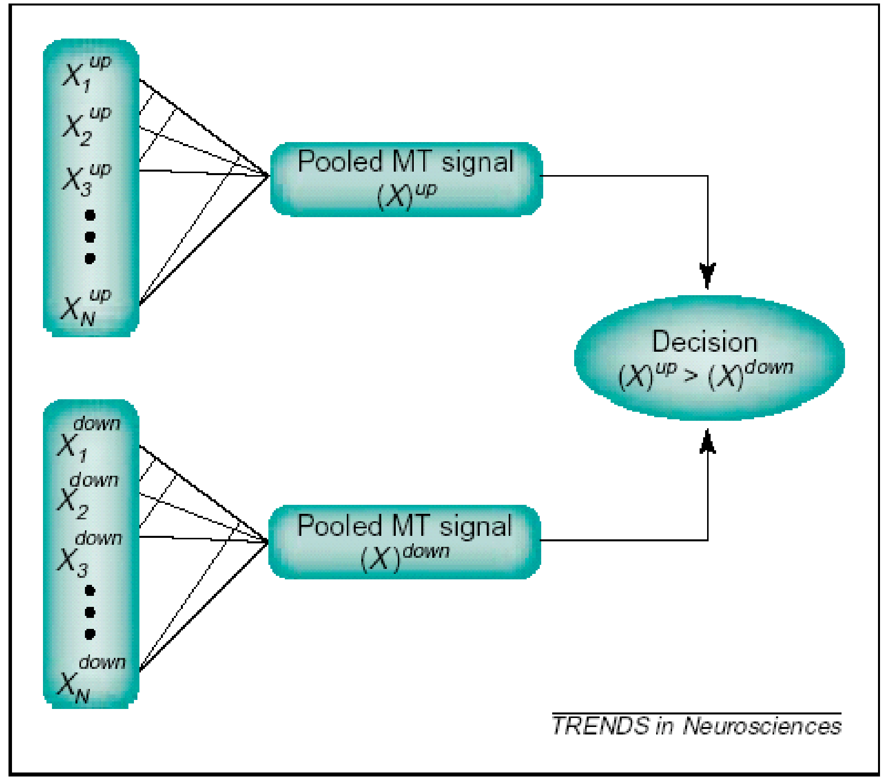
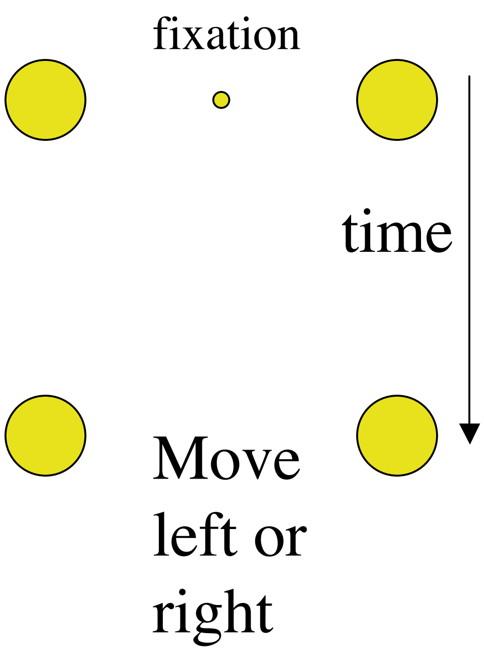


Fig. 4. Outline of the visual-saccadic decision-making model proposed by Shadlen and his colleagues. On each trial, the responses of *N* neurons responding to upwards visual stimulus motion and *N* neurons responding to downwards motion are pooled. Intersecting lines indicate that the responses of neurons responding to the same direction of motion are weakly correlated. Average responses are compared and the larger signal elicits the movement required by the task for that direction of stimulus motion. Reproduced, with permission, from Ref. 22.

where to look[18–21]. In these experiments, monkeys viewed a display of chaotically moving spots of light in which, on any given trial, a subset of the spots moved



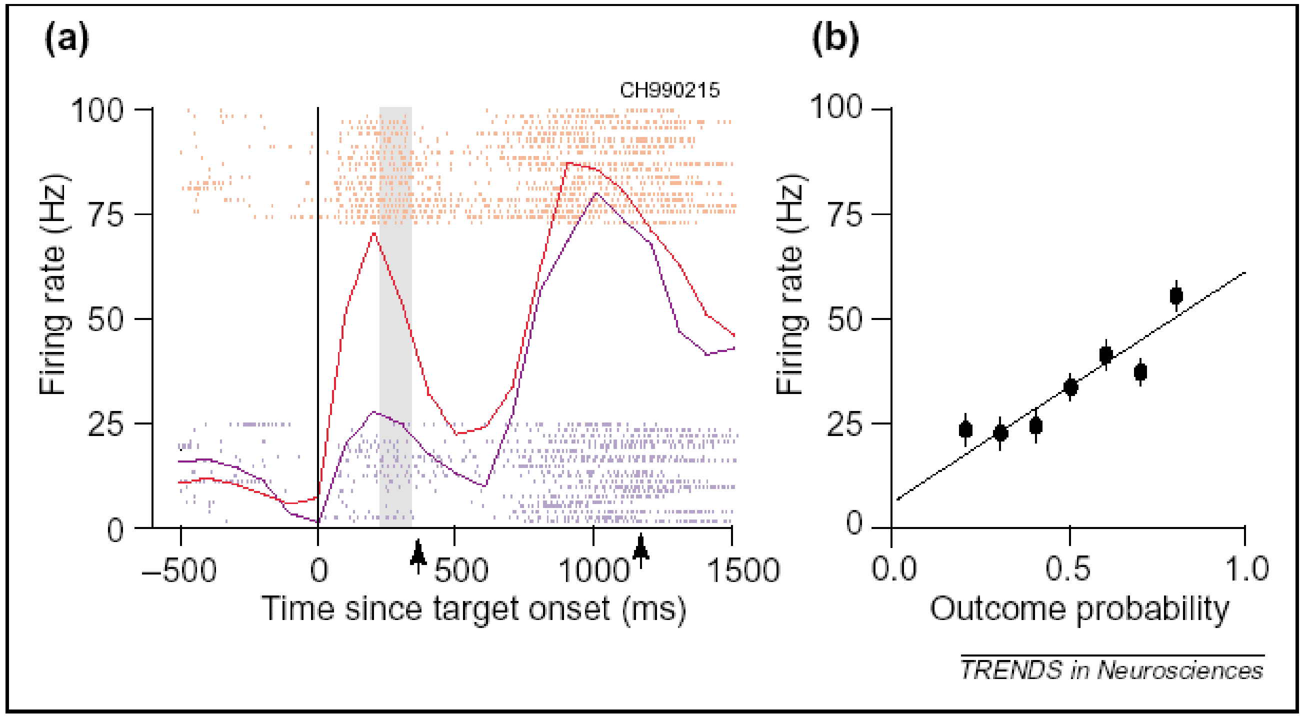


Fig. 5. The probability that a monkey will be instructed to make a particular movement modulates the neuronal activity associated with that movement in the lateral intraparietal area. (a) Average firing rate of an intraparietal neuron on trials that all elicited the best movement for the neuron under study. The red line plots trials on which the best movement was instructed with an 80% probability. The blue line plots trials on which the best movement was instructed with a 20% probability. Raster panels show spike times during the first 20 trials of each type. Black arrows indicate, from left to right, average time of the cue that indicated which movement would actually be required on that trial, and average time of saccade onset. (b) Mean firing rate for the same neuron after stimulus onset. Graph plots neuronal firing rate against the seven different prior probabilities studied in this neuron. Gray bar in (a) shows approximate time of the measured interval. Adapted, with permission, from Ref. 27.