Towards Interaction Design with Active Inference: A Case Study on Noisy Ordinal Selection

Sebastian Stein, John H. Williamson, and Roderick Murray-Smith

School of Computing Science, University of Glasgow, Glasgow, Scotland, UK

Abstract. This paper explores active inference for user interfaces. We implement an active inference approach for 1-of-N selection, a fundamental building block of interactive systems. In this setup, users provide noisy discrete inputs and the interface sequentially identifies an intended target. This problem has an optimal solution (Horstein's algorithm) where the channel noise is iid and known a priori, but is an open problem where the noise is unknown or varying. We reformulate the problem as free energy minimisation and derive a practical active inference implementation. Active inference with a flat noise prior performs comparably to Horstein with conservative noise assumption in the first interaction sequence and as well as Horstein with perfectly calibrated noise thereafter, demonstrating fast adaptation. We also show that active inference can infer the input polarity, offering an extra degree of freedom to users, and adapt to non-stationary noise. The application of active inference to interaction is novel, and we hope this example establishes the groundwork for the community to explore active inference in human-computer interaction.

Code available at https://github.com/drsstein/iwai2024

Keywords: Active Inference · Computational Interaction Design · Ordinal Selection · Adaptive Interfaces · Human Computer Interaction.

1 Introduction

Human computer interaction design is increasingly informed and powered by computational methods: **computational interaction** [12]. Computational interaction optimises interfaces with respect to forward models of user behavior (perception, cognition, motor control, sensor characteristics). This often involves optimisation [2], simulation [11,6], Bayesian inference [17] or reinforcement learning [3,16]. Active inference combines many of the appealing aspects of Bayesian models in interaction design with reasoning over actions that are conventionally approached via RL, but active inference has yet to be applied in a human-computer interaction design context.

Active inference is a model-based control method that minimises expected free energy, simultaneously reasoning about the latent state of its environment and acting to shift the environment state towards some preferred distribution.

Bestowing interfaces with active inference enables them to reason about and adapt to their users and the environment within which they are embedded, while the interaction is ongoing.

Active inference is typically applied to model the behaviour of an agent like a biological system. Here, we characterise the interface as an agent acting under active inference principles whose goal is to extract intention from a user in the face of an unknown and corrupting environment. The agent's goal is to minimise the entropy over the intentional states of the user; it prefers actions that maximise the flow of information from user to system. This is somewhat unfamiliar as an active inference formulation and care is needed in terminology: the agent's *pragmatic goal* is to extract information about user intention and forward it on, but it must also *gain information* about latent states of the user and environment configuration to do so efficiently; the agent wishes to minimise its future surprise about the user's intent. The agent can perceive user actions through sensing, and can act upon the user through the display. This formulation (Figure 1) is a general and powerful way to cast interaction problems as active inference problems.

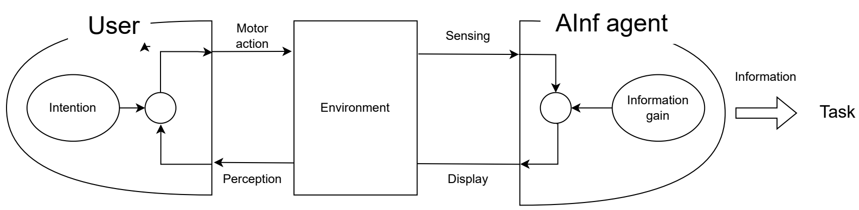

Fig. 1: An active inference agent whose goal is to extract intention from a user's mind and pass it to some external task. The agent has sensing to determine environment states (some of which are influenced by the user) and display channels to influence the environment (some proportion of which are perceived by the user). The agent acts to reduce its uncertainty over user intention so that it can propagate this to the task.

In this paper, we explore the potential for active inference in computational interaction design for noisy ordinal target selection. This is essentially a game where one player (the user) has a number in their mind, and the other player (the computer) tries to work out which number is being thought of by choosing a candidate number and asking the user if the number is higher or lower – with the twist that sometimes the first player lies (noise). Many common interactions can be reduced to 1-of-N selection problems of this nature, either using natural ordering (e.g. alphabets) or imposing an arbitrary order (e.g. a drop-down menu).

2 Related Work

Bayesian models have been explored in the human-computer interaction literature, including at design time (via Bayesian optimisation), interaction time (e.g. via expected information gain) and at evaluation time (Bayesian statistical analyses). Williamson et al. [19] reviews Bayesian models in human-computer interaction generally. With regard to Bayesian approaches to interaction time selection problems, Dasher [18] solved the *entropy-coding problem* of communicating a specific text sequence with a minimum number of inputs, using a feedback model applying arithmetic coding of a statistical language model embedded in a zooming-based interface. [7] used the idea of Bayesian information gain to build a map navigation interface that again minimised the number of inputs required to localise a spatial target. BIGNav formulated the interaction problem as an agent running optimal experiment upon a user. Both Dasher and BIGNav assumed negligible noise in the user inputs.

The *channel coding* problem was explored by Williamson et al. in [20], who proposed the binary selection model explored in this paper and demonstrated robust interface designs for brain-computer interfaces that incorporate Horstein's posterior matching scheme [5]. This results in very robust interfaces in noisy contexts, but is only fully effective for input with known, stationary, independent and identically distributed Bernoulli noise. Simulation as a tool for exploring, optimising and analysing human-computer interaction designs is surveyed in [11], who identify the formulation of explicit *generative* computational models of user (and system) behaviour as a critical step in advancing HCI research. In the active inference literature, [15] explores trust in human–robot collaboration, cast as "mutual predictability" in an active inference framework. [8] describe an active inference approach to an adaptive P300-based text entry BCI, which greatly improved bitrates in communication over a very low capacity speller interfaces. [1] provides a brief outline of the potential use of active inference in brain computer interfaces with the active inference agent engaged in minimising surprise over user intention, following the same line of argument as in this paper. Grizou [4] discuss machine learning approaches to *self-calibrating* interfaces, where the labelling of controls can be arbitrarily permuted but control can still be established; as we will see, self-calibration can arise naturally from an active inference formulation.

3 Problem

The general problem we are interested in is the reliable 1-from-N selection problem (Figure 2), using a binary input subject to random corruption. For example, to pick a number, select a letter, or identify a menu item by pressing one of two buttons. Accumulating several binary button presses is typically required to resolve a particular item. We are interested in *reliable* selection, where users can select items with *arbitrarily low probability of mis-selection* at the cost of an increasing number of binary inputs to identify the target item.

In this abstract setting we ignore the wall clock time to produce each binary decision. This includes the time for the user to reason about the correct decision, the time to actuate the decision and the time to process the feedback from the system about the new state.

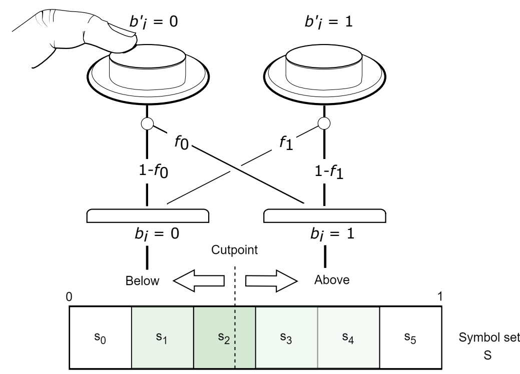

Fig. 2: The ordinal selection problem with noisy binary inputs (adapted from [20]). The user can emit one of two symbols, each with its own probability of flipping.Ordered target symbols are arranged on an interval. A cutpoint m_i is presented to elicit a binary "above/below" decision. A probability distribution over targets (green shading) is maintained during selection.

Although this is an abstract interface model amenable to mathematical modelling, such limited interfaces are practically relevant in brain-computer interfaces such motor imagery EEG control [13] and in other assistive technology input devices such as electromyography where the sensing interface emits binary symbols with very high error levels but where high-capacity, reliable displays are available. The feedback channel is essential; although feedforward error correction (FEC) is commonplace in digital communication, FEC codes are wholly impractical for human input. Feedback error correction is asymptotically more efficient but critically also tractable for human users. We focus on the binary problem for convenience of analysis, but the ideas generalise to any q-ary discrete input device.

3.1 Ordinal selection with noise

Formally, we model an ordinal selection problem, in which the user's task is to identify a discrete symbol $s \in S$ from an ordered sequence of symbols S where typically $|S| = 2^k$. We the user emits a sequence $b_0, b_1, \ldots b_t$ of binary symbols $b_i \in \{0, 1\}$. Timing is ignored and users have a forced choice of binary symbol at each step. There is no explicit selection input like a "click"; decoding is performed once the intended symbol is sufficiently well identified. Each symbol has an associated random flip probability f_0, f_1, f_0 giving the probability that b_i is flipped from 0 to 1 and vice versa for f_1. We assume iid Bernoulli noise. f_0, f_1 fully characterise the communication channel and are referred to as the **channel statistics**. We can model both symmetric iid noise ($f_0 = f_1$) and asymmetric (biased) noise – particularly relevant in brain-computer interfaces where one input symbol is often significantly more corrupted than the other. Figure 2 illustrates our model.

3.2 Known channel statistics: Horstein decoding

In the case where f_0 and f_1 are known in advance *and* there is an effectively noise-free feedback channel *and* we assume that the symbol set is arbitrarily large, $|S| \to \infty$, then Horstein's posterior matching feedback algorithm [5] is known to be optimal. This algorithm operates on an interval of the real number line $0 \leq x \leq 1$ and forms a probability density $f_X(x)$ over possible values of x. The density $f_X(x)$ is represented as a piecewise linear CDF, and the algorithm progresses by adaptively proposing a series of cutpoints $m_0, m_1, ..., m_t$ based on the history of inputs. Each input rescales the CDF about the proposed cutpoint, giving a closed-form update for the posterior density at each step. Each step of the algorithm is equivalent to a Bayesian update, taking the previous PDF as the distribution over the interval and then updating based on the observed evidence. The cutpoints m_i are simply selected at the median density, which is trivial to evaluate from the piecewise linear CDF.

To convert this to an ordinal selection problem, the interval is subdivided into subintervals $[s_l, s_h]$ each corresponding to an element of S such that the unit interval is completely covered by non-overlapping subintervals. The widths of each element of interval are set to the prior probability of each symbol s_i (with uniform width of size $1/|S|$ if a uniform prior over symbols is assumed). The algorithm terminates either when the density concentrates sufficiently in an interval corresponding to a particular symbol $\int_{s_l}^{s_h} f_X(x) > p_k$ with a simple fixed threshold p_k, or alternatively when the (differential) entropy $H(X)$ decreases by some threshold h_k. To allow for a small degree of mismatch between the true channel statistics and the estimated statistics, a *headroom* f_h is added to f_0, f_1 to product f'_0, f'_1; typically f_h is in the range 0.01-0.05. Although the Horstein algorithm is optimal over the continuous space, discretising into a symbol sequence S means that capacity is reduced below the Shannon limit. [20] describes the details of the algorithm and the setting of thresholds.

The known channel statistics problem is solved by Horstein's algorithm modulo details of a particular interface design. However, if f_0, f_1 are *unknown* or changing, or the noise is not iid, the optimal interface for reliable ordinal selection is an open question. Horstein's algorithm is also only optimal for large $|S|$, but many selection problems are from small symbol sets.

4 Method

We assume a model where we have (potentially time varying) probability distributions over the channel statistics $P_t(F_0), P_t(F_1)$ defined by density functions $f_{F_{0,t}}(x)$ and $f_{F_{1,t}}(x)$. Our goal is still to estimate the specific symbol s_i that represents a user's intention. In the unknown channel statistics case, however, we have to trade-off acquiring information about the channel statistics (*channel probes*) and resolving the target symbol. A naïve approach might interleave symbol selection with calibration phases selecting dummy (system-chosen) targets as channel probes. This is simple but inefficient.

We instead choose to formulate this as an active inference problem. We model an agent (AInf agent) whose goal is to perform optimal mind-reading via an unknown channel – to identify the s_i that represents the user's intention with arbitrarily low error rate and in the minimum number of input binary symbols; this implicitly requires online estimation of f_0, f_1. AInf agent must trade off exploration to estimate channel statistics and refine its model of the environment against exploiting the information from the user to identify the hidden target.

Our model formulates the problem as inferring a belief distribution $Q_i(x_i)$ over the unobserved states $x_i : [s^i, f_0^i, f_1^i]$, the unknown symbol s and channel statistics f_0, f_1 at timestep i. The agent's observation at each time step i is the binary symbol $o_i = b_i \in \{0, 1\}$. The agent's only action is to select a cutpoint $m_i \in \mathbb{R}, 0 \leq m_i \leq 1$. The unit interval is mapped onto the ordinal symbols S as described above, using a uniform division of the interval into symbol subintervals, and the symbol is presented to the user. In each step i the agent acquires exactly one binary observation o_i and produces one action m_i.

The belief is initialised with a prior $Q_0(x_0)$ and updated using Bayesian inference. A probabilistic forward model of the environment dynamics $P(x_i|x_{i-1}, m_{i-1})$ is used to propagate the belief through time after every agent action m_i as in Equation (1). A probabilistic forward model of the sensor states $P(b_i|x_i, m_{i-1})$ is used to revise this belief after a new user input is observed as in Equation (2).

$$Q_{i-1}(x_i) = \int P(x_i \mid x_{i-1}, m_{i-1})\, Q_{i-1}(x_{i-1})\, \mathrm{d}x_{i-1} \tag{1}$$

$$Q_i(x_i) = \frac{P(b_i \mid x_i, m_{i-1})\, Q_{i-1}(x_i)}{P(b_i \mid m_{i-1})} \tag{2}$$

$$P(b_i \mid m_{i-1}) = \int P(b_i \mid x_i, m_{i-1})\, Q_{i-1}(x_i)\, \mathrm{d}x_i \tag{3}$$

AInf agent selects actions m_i by minimising the expected free energy $G(\pi)$ over policies $\pi : (m_i, \ldots, m_{i+T-1})$ with time horizon T and choosing the first action of the optimal policy (4).

$$G(\pi) = \frac{1}{T} \sum_{k=i}^{i+T-1} - \underbrace{\mathbb{E}_{Q_k}\left[D_{\mathrm{KL}}\left(Q_k(x_k) \| Q_{k-1}(x_k)\right)\right]}_{\text{Information gain}} - \underbrace{\mathbb{E}_{Q_k}\left[\ln P_k^c(x_k)\right]}_{\text{Pragmatic value}}, \tag{4}$$

where D_{KL} is the KL-divergence and P^c represent the agent's preferences. In this scenario, the agents goal of identifying the the user's intent is aligned with maximising information gain over the full belief state; in a more complex scenario P^c could encourage specific user behavior akin to *nudging*.

5 Implementation

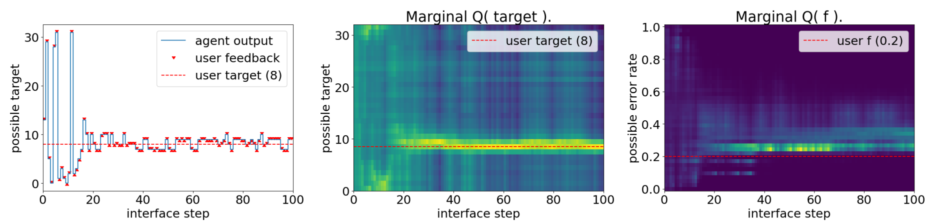

Fig. 3: Interaction between a simulated user with symmetric input error ($f_0 = f_1 = 0.2$) and AInf agent with flat symmetric prior assuming no knowledge about input polarity $Q_0(f_0) = \mathcal{U}(0, 1), Q(f_1) = Q(f_0)$. Left: Sequence of agent actions (blue), user input (red arrows), and user target (red dashed). Center: Marginal belief over user targets. Right: Marginal belief over channel statistics. Note that the belief over user target and channel statistics is in the continuous domain and is only discretized for visualisation purposes.

Commonly, the belief Q is approximated using a Gaussian distribution, which leads to poor approximations of the posterior in this problem. A sequence of user inputs is usually consistent with distant mutually exclusive beliefs. For example, successive "below" inputs are consistent with a low target and low error rate, and simultaneously consistent with a high target and high error rate. To model multi-modal distributions, we use a particle filter. A set of $n_p = 100 \cdot |S|$ particles $\{(x_j, w_j)\}$ represent the belief as a mixture of Dirac delta distributions (see Equation (5)), where $w_j \in \mathbb{R}^+$ represent normalised weights $\sum_j w_j = 1$. Particle weights are updated using (2) after every observation, and particles are resampled when the effective sample size $1/\sum w_j^2$ falls below a threshold $\tau_w = 0.5$ using low variance re-sampling. The temporal dynamics of the system are assumed to be stationary with some diffusion $\sigma_d = 0.001$ on the error rate to accommodate drift, updating x_j by sampling from (6).

$$Q(x) = \sum_j w_j \, \delta(x_j) \tag{5}$$

$$P(x_{i+1} \,|\, x_i, m_i) = \mathcal{N}(x_i, \sigma_d^2) \tag{6}$$

Here, the set of possible actions has finite size $|S|$. We exhaustively evaluate all $|S|^T$ policies with time horizon $T = 1$. Where unobserved states are independent of the agent's actions, control reduces to solving a Bandit problem. There, longer time horizons do not improve action selection. The D_{KL}-divergence in (4) is computed on the Bernoulli distributions w_i and w_{i-1}. Figure 3 shows an example interaction between a simulated user and the agent.

6 Evaluation

We evaluate the performance of the active inference agent through computational simulation experiments designed to answer the following research questions:

- RQ1: How quickly can AInf agent infer channel statistics and user targets?
- RQ2: How do AInf agent decisions compare to Horstein's algorithm?
- RQ3: How well does AInf agent adapt to non-stationary channel statistics?

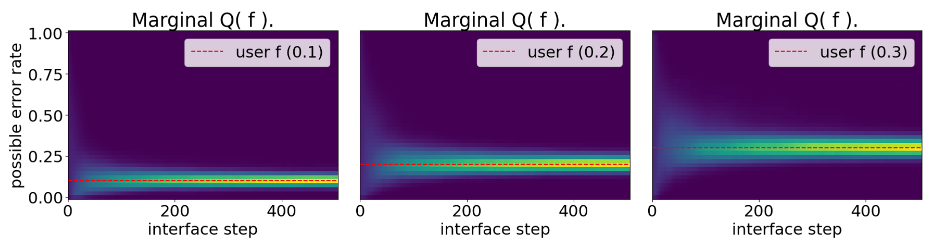

Fig. 4: Marginal probability of the true channel statistics under the belief distribution. Mean marginal belief distribution across all 32 user targets and 10 repetitions each over five successive interaction episodes of 100 steps. The belief converges around the true channel statistics within 100 to 300 steps, with higher error rates taking longer to infer.

RQ1: How quickly can AInf agent infer channel statistics and user targets? We simulated users with symmetric channel statistics $f_0 = f_1 \in [0.1, 0.2, 0.3]$ and every symbol as target $0 \leq s < |S| = 32$, each interacting with AInf agent 10 times for 500 steps resulting in $3 \times 32 \times 10 = 960$ simulation runs. AInf agent belief was initialised assuming error symmetry $f_0 = f_1$ and *without* assuming knowledge about input polarity $Q_0(f) = \mathcal{U}(0, 1)$. Within each run, the agent's belief over user targets was reset to a flat prior every 100 steps, while the belief over channel statistics was maintained, simulating five sequential episodic interactions. We measured the probability of true error $f^* \pm \epsilon$ under the belief distribution at every step in each run, shown in aggregate in Figure 4. Starting from a flat prior, the belief about the true channel statistics increases over time demonstrating that they are being identified. Smaller error rates are identified faster than higher error rates. Under RQ3 below we show that AInf agent target selection behavior is comparable to that under perfect knowledge of channel statistics after 1-3 episodes.

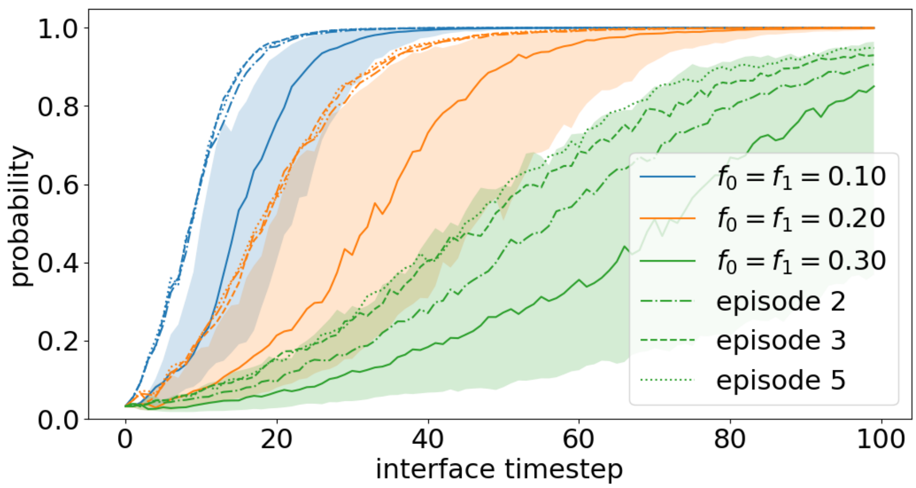

Fig. 5: Marginal probability of the true user target under the belief distribution. Shown are the median and inter-quartile range across all 32 user targets and 10 repetitions in the first 100 interaction steps, and the median across successive interaction episodes. Target inference speed converges after 1-3 episodes demonstrating the saturating effect of inferring the channel statistics.

We evaluate user target inference on the same set of simulations, measuring the probability of the user's target under the belief distribution in the first and subsequent episodes (Figure 5). User targets are inferred faster when error rates are lower, and the number of steps required to infer the user target converges after 1-2 episodes, demonstrating the saturating benefit of inferring channel statistics on target inference.

RQ2: How does AInf agent decision-making compare to the Horstein's algorithm? We evaluated decision-making performance using speed-accuracy tradeoff curves. We measured the decision accuracy (higher is better) and the number of interaction steps until a decision was taken (lower is better) under a decision rule that selects the most likely symbol $\operatorname{argmax} Q(s)$ under the belief distribution when $\max Q(s) \geq \tau$ for varying thresholds $\tau \in [0, 1]$.

AInf agent decision-making performance was evaluated on the same simulations as in RQ1 above. We repeated these simulations with different priors over channel statistics $Q_0(f) = \mathcal{U}(0, 0.5)$. We shall refer to this set of simulations as *known control polarity* and to the original set of simulations as *unknown control polarity*, as error rates greater than 0.5 can be interpreted as predominantly, possibly intentionally, sending the flipped signal. This problem can arises in several interaction contexts. For example, joystick motion "away from" and "towards" the user's body is mapped differently onto "up" and "down" control actions in aeroplane and helicopter control. Inferring and adapting to polarity in the user's mental model of motion-to-control mappings offers an extra degree of freedom to

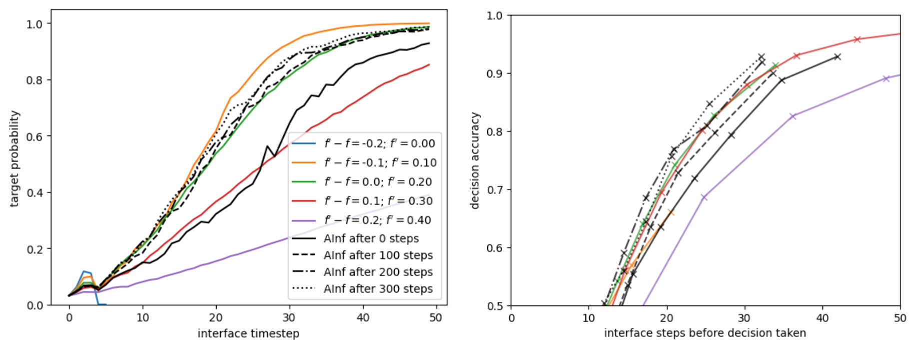

Fig. 6: Decision-making performance of AInf agent and Horstein's algorithm with *known control polarity*. Median marginal belief probability of the true user target for $f_0 = f_1 = 0.2$, varied assumed headroom $f' - f$ for the Horstein algorithm, and sequential episodes for the AInf agent.

users [14]. To evaluate decision-making performance of Horstein's algorithm, we simulated interactions of users with symmetric channel statistics $f_0 = f_1 = 0.2$ and every symbol as target $0 \leq s < |S| = 32$, each interacting with a variant of Horstein's algorithm 100 times for 100 steps. Ten variants of Horstein's algorithm with different *headroom* $(f' - f) \in [-0.2, -0.1, 0, 0.1, \ldots 0.8]$ were used, resulting in $32 \times 100 \times 10 = 32000$ simulation runs. In every run, the belief over user targets was initialised to a flat prior.

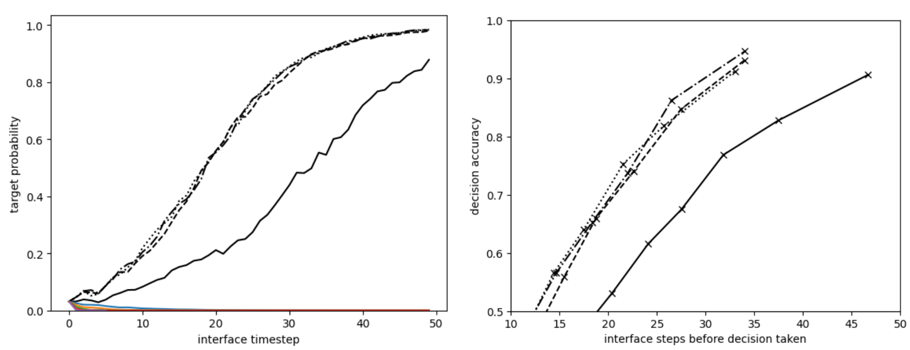

Fig. 7: Decision-making performance of AInf agent and Horstein's algorithm with *unknown control polarity*. Median marginal belief probability of the true user target for $f_0 = f_1 = 0.2$, varied assumed headroom $f' - f$ for the Horstein algorithm, and sequential episodes for the AInf agent.

The target probability over time and speed-accuracy curves are shown for the *known control polarity* and *unknown control polarity* conditions in Figures 6 and

7, respectively. In the *known control polarity* condition, AInf agent target inference speed and decision-making performance are comparable to Horstein's algorithm with perfect knowledge of the channel statistics after 200 steps, and outperforms it slightly thereafter. This demonstrates that inferring channel statistics online from a flat prior is feasible, and that AInf agent can address the reliable ordinal selection. Horstein's algorithm decision-making performance appeared very sensitive to its *headroom*. Horstein's algorithm with negative headroom, underestimating error rates, and headroom above 50%, i.e. mismatched polarity, failing catastrophically (speed-accuracy curves are near zero accuracy across all timesteps. Where the headroom was too conservative ($f' - f = 0.2$), overestimating error rates, accuracy was around 10% lower at the same number of interaction steps compared to the best performing Horstein variant.

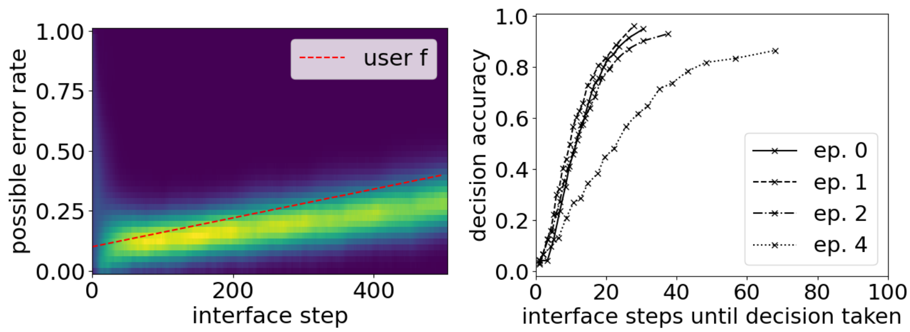

Fig. 8: Marginal belief distribution over channel statistics (left) and decision-making performance (right) of AInf agent with *unknown control polarity* and non-stationary channel statistics.

RQ3: How well does AInf agent adapt to non-stationary channel statistics? Above, we have shown that AInf agent can infer and adapt to unknown channel statistics assuming those statistics are stationary. To explore how well AInf agent responds to smooth changes in the channel statistics, we performed a series of simulations as in the *unknown control polarity* condition, but with symmetric error rate increasing linearly from 0.1 to 0.3 over 5 successive episodes of 100 timesteps. We increased the diffusion scale to $\sigma_d = 0.01$ to make such changes more likely under the agent's model. We evaluated how well AInf agent can track changing channel statistics by estimating the mean marginal belief distribution over error rates across timesteps, and the decision-making performance via the speed-accuracy curve (see Figure 8). While the agent is more uncertain about the channel statistics overall, it continues to provide a good speed-accuracy trade-off throughout the period of error rate degradation.

In preliminary experiments not shown here for lack of space, we found AInf agent decision-making performance to remain stable under a wide range of particle filter hyperparameter settings, including changes to the number of particles n_p, resampling threshold τ_w and amount diffusion scale σ_d.

7 Discussion

We presented an active inference approach to reliable selection with two noisy inputs. This is human-computer interaction stripped back to its barest elements, but still complex enough to represent real interactive systems. We formulate the interface as an independent agent charged with facilitating the flow of information, acting as an active transducer able to reason about the environment and user characteristics to optimise this flow. Active inference gives a high-level formulation of the problem that is fully Bayesian and flexible enough to precisely model this task. Classical information theoretic approaches are only optimal under assumptions that are rarely met in the messiness of human interactions.

We focused on the control polarity and non-stationary channel statistics, but it would be quite feasible to relax other assumptions in the active inference formulation. These include the iid noise assumption (for example, modelling bursty or otherwise correlated noise), time-varying numbers of input symbols or input symbols with different costs. This scenario often arises in assistive technology [10] where some inputs may have high physical demands or long refractory periods.

Approximate Bayesian inference using a particle filter is well suited to modelling multi-modal distributions but has a high computational cost. As we move from the most elementary interactive systems explored in this paper, we will need to evaluate these more complex tasks with real users. To achieve real-time (sub-second) closed-loop performance with an active inference approach we may need to implement amortized inference approaches, such as proposed in [9].

In the ordinal selection task, stationarity in the unobserved state and alignment of the agent's goal with maximising information gain simplified the inference problem, but time-varying dynamics and exploration-exploitation trade-offs abound in human-computer interaction tasks. Relying only on forward models of users and applications, active inference can help make offline and online interaction design more transparent and modular. The unifying approach of active inference holds real promise in human-computer interaction, bringing together Bayesian models of interaction with explicit reasoning over future actions. While computational demands and challenging modelling work lie ahead, as we have demonstrated in this paper, even simple and well-understood interaction problems can be turbocharged by an active inference perspective.

Acknowledgements

This research was supported by the project Designing Interaction Freedom via Active Inference (DIFAI). The project was selected by the ERC (Advanced Grant proposal 101097708) and was funded by the UK Horizon guarantee scheme as Engineering and Physical Sciences Research Council (EPSRC) project EP/Y029178/1. We thank Aegean Airlines S.A. for their donation.

References

1. Ather, S.H.: Active inference as a framework for brain-computer interfaces. bioRxiv pp. 2021–02 (2021)
2. Dayama, N.R., Todi, K., Saarelainen, T., Oulasvirta, A.: GRIDS: Interactive Layout Design with Integer Programming. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–13. ACM, Honolulu HI USA (Apr 2020). https://doi.org/10.1145/3313831.3376553
3. Glowacka, D., Howes, A., Jokinen, J.P., Oulasvirta, A., Şimşek, Ö.: RL4HCI: Reinforcement Learning for Humans, Computers, and Interaction. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–3. ACM, Yokohama Japan (May 2021). https://doi.org/10.1145/3411763.3441323
4. Grizou, J.: Interactive introduction to self-calibrating interfaces (Dec 2022). https://doi.org/10.48550/arXiv.2212.05766
5. Horstein, M.: Sequential transmission using noiseless feedback. IEEE Transactions on Information Theory **9**(3), 136–143 (1963)
6. Ikkala, A., Fischer, F., Klar, M., Bachinski, M., Fleig, A., Howes, A., Hämäläinen, P., Müller, J., Murray-Smith, R., Oulasvirta, A.: Breathing Life Into Biomechanical User Models. In: Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology. pp. 1–14. ACM, Bend OR USA (Oct 2022). https://doi.org/10.1145/3526113.3545689
7. Liu, W., D'Oliveira, R.L., Beaudouin-Lafon, M., Rioul, O.: BIGnav: Bayesian Information Gain for Guiding Multiscale Navigation. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 5869–5880. ACM, Denver Colorado USA (May 2017). https://doi.org/10.1145/3025453.3025524
8. Mladenovic, J., Frey, J., Joffily, M., Maby, E., Lotte, F., Mattout, J.: Active Inference as a unifying, generic and adaptive framework for a P300-based BCI. Journal of Neural Engineering **17**(1), 016054 (2020)
9. Moon, H.S., Liao, Y.C., Li, C., Lee, B., Oulasvirta, A.: Real-time 3d target inference via biomechanical simulation. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. pp. 1–18 (2024)
10. Müller-Putz, G.R., Ofner, P., Schwarz, A., Pereira, J., Luzhnica, G., di Sciascio, C., Veas, E., Stein, S., Williamson, J., Murray-Smith, R., et al.: Moregrasp: Restoration of upper limb function in individuals with high spinal cord injury by multimodal neuroprostheses for interaction in daily activities (2017)
11. Murray-Smith, R., Oulasvirta, A., Howes, A., Müller, J., Ikkala, A., Bachinski, M., Fleig, A., Fischer, F., Klar, M.: What simulation can do for hci research. Interactions **29**(6), 48–53 (2022)
12. Oulasvirta, A., Kristensson, P.O., Bi, X., Howes, A.: Computational Interaction. Oxford University Press (2018)
13. Padfield, N., Zabalza, J., Zhao, H., Masero, V., Ren, J.: EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges. Sensors (Basel, Switzerland) **19**(6), 1423 (Mar 2019). https://doi.org/10.3390/s19061423
14. Reddy, S., Levine, S., Dragan, A.D.: First contact: Unsupervised human-machine co-adaptation via mutual information maximization. arXiv preprint arXiv:2205.12381 (2022)
15. Schoeller, F., Miller, M., Salomon, R., Friston, K.J.: Trust as extended control: Human-machine interactions as active inference. Frontiers in Systems Neuroscience p. 93 (2021)

16. Todi, K., Bailly, G., Leiva, L., Oulasvirta, A.: Adapting user interfaces with model-based reinforcement learning. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2021)
17. Velloso, E., Morimoto, C.H.: A Probabilistic Interpretation of Motion Correlation Selection Techniques. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–13. ACM, Yokohama Japan (May 2021). https://doi.org/10.1145/3411764.3445184
18. Ward, D.J., Blackwell, A.F., MacKay, D.J.C.: Dasher—a data entry interface using continuous gestures and language models. In: Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology - UIST '00. pp. 129–137. ACM Press, San Diego, California, United States (2000). https://doi.org/10.1145/354401.354427
19. Williamson, J.H.: An Introduction to Bayesian Methods for Interaction Design. Bayesian Methods for Interaction and Design **1** (2022)
20. Williamson, J.H., Quek, M., Popescu, I., Ramsay, A., Murray-Smith, R.: Efficient human-machine control with asymmetric marginal reliability input devices. Plos one **15**(6), e0233603 (2020)