AUTOMATION RELIABILITY AND DECISION STRATEGY: A SEQUENTIAL DECISION MAKING MODEL FOR AUTOMATION INTERACTION

Aditya Acharya, School of Computer Science, University of Birmingham, Birmingham. UK
Andrew Howes, School of Computer Science, University of Birmingham, Birmingham. UK
Chris Baber, School of Engineering, University of Birmingham, Birmingham. UK
Tom Marshall, Institute of Applied Healthcare Research, University of Birmingham, Birmingham. UK

The question of how people make use of automation to support their decision making is becoming increasingly important. As computers provide ever greater input to the collection, analysis and interpretation of data, so they are more likely to be partners in decision making. However, when automation makes recommendations that the human disagrees with or that might be based on erroneous analysis, then this could result in a change in decision strategy. It is not simply a matter of ignoring or rejecting the recommendation but rather a matter of deciding how best to make use of the automation's output. By modeling information search and decision strategies under different levels of information reliability, we demonstrate that it makes sense to adapt decision strategy to the information context.

Introduction

From their meta-review of studies reporting human interaction with automation, Wickens and Dixon (2007) proposed that there is a cut-off at around 70% reliability beyond which human performance is worse with automation than without. But, as Lee and See (2004) note, "*The absolute level of the drop off seems to be highly system and context dependent*" (p.72). Wickens and Dixon (2007) proposed that (from the studies that they reviewed) *"…operators chose to depend on the imperfect automation, knowing that it is far from perfect, in order to preserve available processing resources for other tasks*." (p.208). It is this question of when operators will choose to rely on imperfect automation, that is of interest in this paper.

One implication of Wickens and Dixon's observations is that the output from automation could be informationally useful, even when it is incorrect. The output could, for example, help frame the question that the operator might wish to ask in order to check the information, or it might provide an anchor against which to judge other data, or it might simply be that the operators are *"…giving those tasks whatever resources are required to preserve constant performance level…"* (p.208). If this is the case, then one might expect a trade-off between the resource costs in acquiring information and the expected value of that information for a decision. Previous work has shown that this trade-off can be understood through user modeling (Chen, Bailly, Brumby, Oulasvirta, & Howes, 2015; Chen, Starke, Baber, & Howes, 2017).

The current article reports a user model of the trade-offs between the costs of acquiring information and the information reliability. The model is developed for a typical decision problem faced by a General Practitioner (GP) (a primary care physician) involved in oncology referral decisions.

Task Domain

For GPs, the decision of whether to refer oncology patients for tests can be challenging. The presentation of symptoms might be unusual for GPs and they might not have detailed experience of diagnosing a specific form of the disease, compared to the more common ailments with which patients typically present. GPs might also be concerned about making too many referrals, and so might adopt a 'conservative' decision policy in which they seek definite and clear information prior to making their decision. However, the available information (in terms of patient reports and symptoms) could be fuzzy, i.e., ambiguous, unclear and incomplete. In the UK, the National Institute for Clinical Excellence (NICE) guidelines for referral relate combinations of symptoms to referral decisions. Indeed, one of the motivations for the work undertaken here is why do some GPs *not* follow the guidelines? We suspected that this related to the GPs confidence in the reliability of the guidelines and the relative 'costs' associated with using (or not using) them. For example, the guidelines might recommend a particularly invasive test in order to determine whether a patient has a given condition. However, the GP might weigh the potential diagnostic benefit of the results of the test against the level of discomfort (physical and psychological) that the patient might experience from the test. So, there is a cost associated with each information source, e.g., Full Blood Count test is a simple blood test and has low cost in terms of patient discomfort and time, but Faecal Occult Blood test requires specialized equipment, is time consuming and uncomfortable. In order to make the right decision, a GP has to take into account the cost involved in the diagnosis process. While, successful referrals are more rewarding for a GP (Hamilton, 2010) , a wrong referral for benign illness adds financial costs to NHS (Laudicella, Walsh, Burns, & Smith, 2016).

Against these concerns, the GP would seek to ensure that tests are applied as early as possible in order to manage conditions. Taken together, these points indicate that each case creates a different trade-off space in which the anticipated cost of an action is weighed against the expected benefits. This means that there might be some cases in which low reliability guidelines could be followed, particularly in highly sensitive medical cases. For the purposes of this paper, we assume that the question of whether or not to follow guidelines (which often present an algorithmic relation between specific symptoms from the patient and the proposed course of action of three GP) can be analogous to the decision as to whether or not to accept recommendations from automation.

Model

We present a model of how a human decision maker integrates their own preferences and knowledge with advice from automation. The model is formalized as a *sequential decision process*. Work on simulating clinical decision making using sequential decision models has previously been reported by Alagoz, Hsu, Schaefer, and Roberts (2010), Bennett and Hauser (2013) and many others.

One widely adopted way to rigorously describe a sequential decision process is as a Markov Decision Process (MDP). We will give a more formal definition of an MDP in a following paragraph, but for now consider the following properties. First, an MDP offers a formal means of defining processes and actions with uncertain outcomes. For example, a clinical test may prove positive or negative and it will do so with greater or lower reliability. Second, an MDP can be used to model how a sequence of multiple small information gathering actions can build evidence for a diagnosis. Further, the decision as to when to stop gathering information can be revisited after each action. In other words, the MDP can be used to model interactive decision making. Third an MDP defines a machine learning problem and a range of well-known algorithms can be used to find efficient diagnostic policies. These might determine, for example, which referrals to gather in which circumstances.

The empirical validity of sequential decision processes in predicting human behavior has previously been shown for a range of related tasks, including menu search (Chen et al., 2015), credit card fraud detection (Chen et al., 2017) and visual search (Acharya et al., 2017). For example, in the fraud detection task the model explains the decision of classifying a transaction fraudulent or not as an optimal adaptation to the user information processing constraint, the validity of each information cue available for diagnosis and the cost of looking up information.

In this work we define the sequential decision process as a Markov Decision Process (MDP) which is a framework to mathematically formulate sequential decision-making problems. Previously, the model has been used to consider decision making in financial fraud domain (anon, 2015). In the MDP framework, interaction with the external environment is assumed to be in discrete time steps. At any time step, t, the agent perceives the current state of the environment, s_t. The agent then has to decide on an action, a_t, to choose that is available in state s_t. On executing the action a_t, the environment then transits to a new state s_{t+1} with a probability defined by the transition function $T(s_t, a_t, s_{t+1})$, and feedback in form of a reward, r_t, is provided to the agent. The reward r_t serves as an indicator to the agent as to how good the action a_t is when taken in state s_t. The goal for the model is to make a sequence of independent action selection choices (until termination) such that it maximizes a discounted cumulative reward over a potentially infinite horizon. With enough trials, the model learns the optimal strategy. In the following sections more detail is provided about how the state and action selection work.

External Display

In the model developed for this paper, 'patient information' is presented under a number of tabs (see Figure 1). In this user interface, information is available on separate pages. To access each page, a tab at the top of a page is clicked. For the model, it is assumed that clicking on a tab would result in the information on that tab being read and understood correctly. It is possible, of course, to set parameters in the model that reflect errors in interpretation, but for this work it is assumed that reading is error-free. The first tab presents patient personal information, for example, patient symptoms, age, BMI and Blood Pressure. The second tab presents the test results, which includes haemoglobin count, white blood cell count, platelets, mean cell volume, ferritin and faecal occult blood result. The third tab shows the automated recommendation, i.e., either to refer or take no further action.

Figure 1: Tabbed information pages

In addition to selecting and reading information pages, the model is also able to ask for further tests (which result in additional information pages to read) or refer the patient (which ends the decision task) or take no further action (which also ends the decision task).Thus, for the model, the set of Actions that could be performed is: Click on Patient Information tab; Click on Blood Count tab; Click on Guidance tab; Ask for FOB test; Ask for Ferritin test; Refer; No further investigation. The 8 information cue presented under the tabs have specified ranges used by the model and interpreted in terms of normal and abnormal readings (table 1). It is worth noting that the distribution of parameters shown in table 1 indicate a broad of values for some very specific conditions. Future versions of the model would use a broader range of parameters and conditions.

Table 1: Cue Distribution

Cue Name	Normal	Abnormal
Faecal Occult Blood	Negative	Positive
Haemoglobin	135 to 160	104 to 130
Age	40 to 60	60 to 70
Ferritin	14 to 24	6 to 14
Body Mass Index	22 to 28	16 to 22
Mean Cell Volume	78 to 91	70 to 78
White Blood Cell	4.0 to 8.0	< 4.0
Platelets	190 to 351	< 190

Each cue also has an associated validity, i.e., probability that the cue indicates an Abnormal reading given that the patient has cancer. While we are interested in the influence of automation reliability of decision making, this has to be considered in context (as noted in the introduction). For this paper, the context is defined as the validity of the information available to the decision maker. In the model three levels of cue validity were explored (figure 2), and one can see that most of the 8 cues have a similar validity (0.5) and two of cues have higher validity.

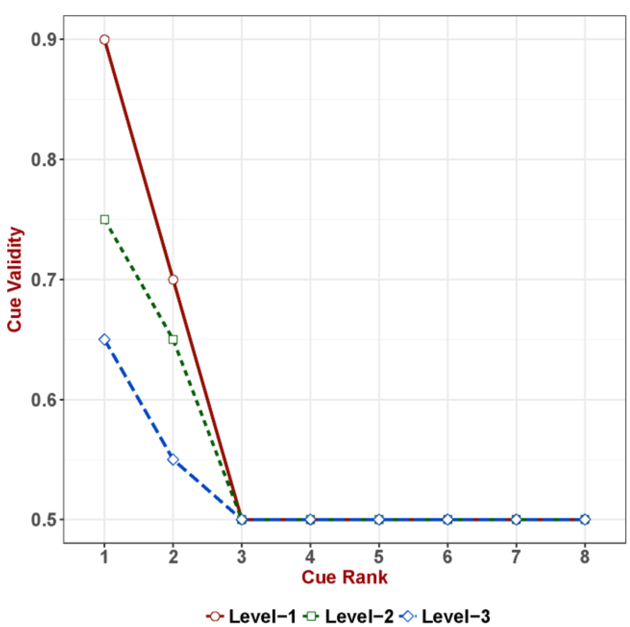

Figure 2: Levels of cue validities used in the model.

We also varied the automated recommendation accuracy over 11 levels [100%; 98%; 96%; 94%; 92%; 90%; 85%; 80%; 75%; 70%; 50%].

Thus, the model seeks to optimize reward (in terms of using higher reliability information sources, either cues or guidance) in terms of the cost of accessing the information.

Actions

As noted above, the action space consists of (1) open information tab, (2) ask for more tests, (3) Refer patient, (4) No further action. In our model, there were three information tabs that the model could open to access patient information, or ask for repeat of full blood count test or the faecal occult blood test. A trial was terminated when the model either choose to refer or take no further action.

Reward

A Time cost was associated with every information gathering action. A scalar penalty of 5s was given for every tab click. This provides a base cost of accessing information. Additional cost also applied if the model made multiple requests for further tests (on the assumption that such behavior indicated over-referral). So, when the model asked for repeats of the full blood count test, a scalar penalty of 5s was given, and for faecal occult blood test a penalty of 10s was given. On successful diagnosis, i.e., to refer when patient has cancer or take no further when patient does not have cancer, a reward of 50-TotalTime was provided. If the model decides to refer and the patient does not have cancer a penalty of 50s was added to Total Time and if no action was taken but the patient had cancer a penalty of 200s was added to Total Time. The cost function enforces a time pressure on the model, where, foraging for information to achieve higher accuracy comes at a cost.

State

In the model a state represents the true underlying information presented on the display. In the model there were 8 information cues and 1 guidance information spread across three different tabs. A state is a 9-element vector. Each element of the vector represents the information gathered for one of the cues. The value of each cue was discretised into three levels, representing 'Abnormal (AB)', 'Normal (N)' and 'Unknown (U)' respectively. This gives the 9-element vector [AB,N,N,AB,AB,AB,N,AB,N]. Hence, the size of the state space is $3^9 = 19683$.

Learning

The model's knowledge is stored in a tabular form that maps each state with corresponding actions. This mapping is learned using a machine learning algorithm called Q-learning (Watkins & Dayan, 1992). Q-learning defines a set of states, *S*, in an environment and a possible set of actions, *A*, in those states. At each iteration it learns the value of each of those actions for each state; this value, *Q(s, a)*, is the state-action value. A walk-through of the algorithm is provided below.

At the beginning, an empty table (Q-table) was assumed representing no knowledge. All values (i.e., Q-values) were initialized to 0. The model was trained by interacting with the environment through trial and error until the performance plateaued (10^6 trials). The model choose what action to take using an e-greedy algorithm, i.e., it choose actions greedily based on the highest *Q(s, a)* value with a probability of 1-ε, otherwise, chooses a random action. The value for ε was set to 0.01 in our model. Q-values were updated on each iteration according to the reward, as shown in equation [1]

[1]

where, *Q(s,a)* is the Q-value for one state-action pair (*s*,*a*), *r* is the immediate reward obtained while the action *a* is taken (based on the reward function defined above); α is the learning rate, which is set to 0.1; γ is the discounted factor, which is set to 0.9. The optimal policy acquired through this training was used to generate the predictions (last 1000 trials of the simulation).

Results

The model was run for each of the cue validity levels (figure 2) and across each of the 11 levels of guidance accuracy. The following graphs show the frequency with which the model followed the guidance under these different conditions. The graphs, therefore, show

In each of the following graphs, the frequency of guidance tab click against varying guidance accuracy for different time cost of looking up guidance as a fraction of mean cost of information. In other words, the assumption is that high cost for accessing guidance could deter the model, unless the guidance was perceived to be highly reliable.

Level 1 Cue Validity: In level 1, two of the Information Cues are 90% and 70% reliable (all other cues were set to 50%). As figure 3 shows, when the cost of the guidance is much lower than other information sources (0.2 times), then the model consults it from around 70% accuracy. When the cost of accessing guidance was the same of the majority of other cues, the model would consult guidance when the reliability was over 85% (figure 3). As the cost of accessing guidance increases (i.e., as the curves on the graph are towards the right), so the frequency of consulting it decreases unless its reliability is around 98%. In this way, the model shows a trade-off between the relative cost of accessing guidance and the reliability of the guidance.

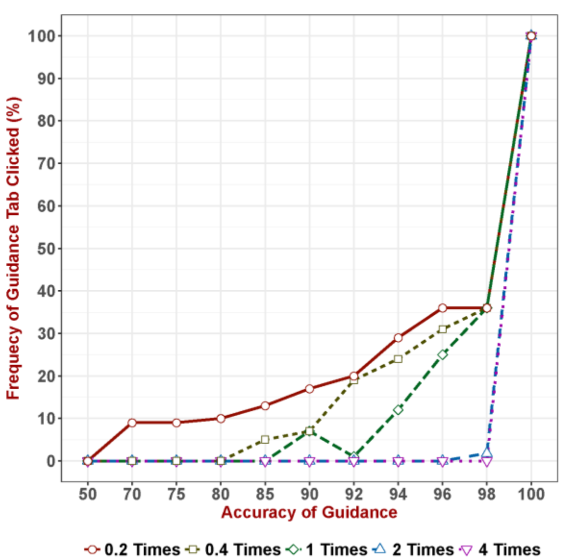

Figure 3: Level 1 cue validity

Level 2 Cue Validity: In this condition, two of the information cues are around 65% and 75% reliable (all other cues were 50%). In this condition, accessing guidance when it has a similar cost to the other cues, begins to rise when reliability exceeds 75% (figure 4). The division between low and high cost of access is much clearer in figure 4 than in figure 3. For the model to consult guidance that is more costly to access than the other cues, this needs to have reliability of over 90%.

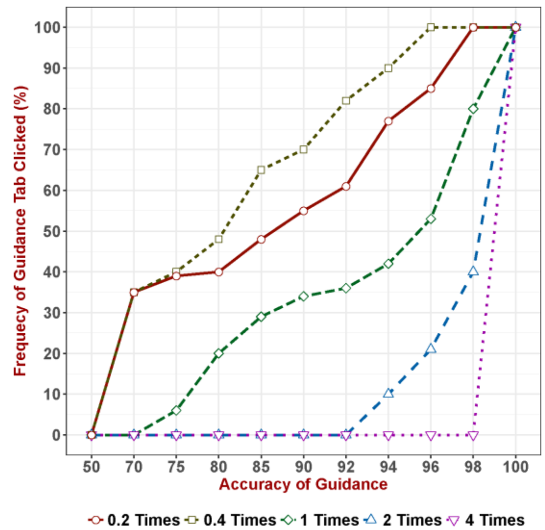

Figure 4: Level 2 cue validity

*Level 3 Cue Validity***:** When the other cues are of 50%, then the model is more likely to consult guidance (figure 5). The threshold here seems to be around 70% (unless the cost of accessing guidance was much lower than accessing the other cues).

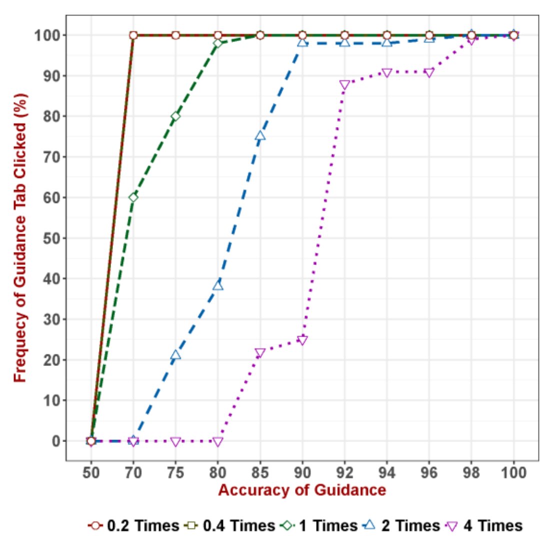

Figure 5: Level 3 cue validity

Discussion

In this paper, we presented a model of how a rational decision maker integrates advice from automation. The results indicate that strategies for information gathering and decision making are an emergent consequence of the reliability of the information sources, relative to that of the guidance, and the cost of accessing these sources. We propose that this defines the ecology of the decision environment and that the results indicate that rational decision making (i.e., a decision maker that is seeking to optimize reward with minimized cost) will trade-off use of guidance relative to its cost and reliability.

Given the review of studies from Wickens and Dixon (2007), it is interesting that we find a 70% threshold in some instances. Not only does this reflect human behavior but also our modeling provides us with some principles that can be used to explain this behavior. First, the reliability of the guidance makes sense only in terms of the decision ecology. In other words, it is not sufficient to define automation reliability in isolation, but rather this needs to be considered in terms of the information context in which it is provided. This echoes points those made by Wickens and Dixon (2007). Second, the decision ecology involves a trade-off between the utility of the information source and the cost of accessing it. In our model, cost was defined by three parameters: access time (such that consulting each new piece of information incurred a cost); a penalty on over-referring (such that the choice of action could not simply involve repeatedly asking for more tests); and a penalty for selecting an incorrect action. One could imagine a revision of the model in which the definition of cost was elaborated, perhaps to include economic cost associated with testing, or emotional cost (to the patient) of having to undertake the test. However, we believe that the basic pattern of results would reflect those presented here.

The decision making problem was formulated as a sequential decision making problem with a machine learning algorithm to find an optimal decision strategy. We propose that this is a novel approach to both the study of human decision making and to the framing of decision problems for Artificial Intelligence. By describing a human decision policy in terms that are tractable to AI, we can explore massive data sets in ways that humans can find understandable. In our previous studies, we have shown that the model predicts the strategies that people will use for information search (Chen et al., 2017). We propose that the model reported in this paper demonstrates how strategies for combining advice from automation with other sources of information will be influenced by the trade-off between the time cost of acquiring information / advice and the predictive accuracy of that information/advice. Ongoing work is being conducted in which the information (shown in figure 1) is presented (on a web-based task) to GPs. This will provide insight into decision time and information choice (which can be compared with the model results). A benefit of running the model is that we are able to contrast many more conditions than would be feasible with even the most enthusiastic of human participants. The model also provides a basis for developing hypothesis for further, e.g., in terms of those versions of the decision ecology which are most likely to prove challenging for the decision maker.

Our results show that guidance does not have to be 100% accurate for it to be useful to decision makers. However, decision makers will consider the guidance relative to the information context in which it is provided. If there are information sources that are perceived to have higher reliability then these would be given preference, even if this might counteract the guidance. Relating this to the question which motivated this particular study, we can say that it would be rational for General Practitioners to ignore NICE guidelines if (a) they felt that there was other information that would necessitate a different decision, and (b) there was perception that the guidelines themselves were no entirely accurate. One consequence of this work is that rather than seeking enforce compliance with guidelines, it might make more sense to calibrate the accuracy of these guidelines (in terms of the relationship between referral decision and specific symptoms). We have access to a large database of clinical decisions and are undertaking that evaluation over the coming months. Our study also shows that lower reliability guidance is only useful when it is combined, under the right conditions, with other sources of information.

References

Acharya, A., Chen, X., Myers, C. W., Lewis, R. L., & Howes, A. (2017) Human Visual Search as a Deep Reinforcement Learning Solution to a POMDP, *Proceedings of the Annual Meeting of the Cognitive Science Society* , 51-56.

Alagoz, O., Hsu, H., Schaefer, A. J., & Roberts, M. S. (2010) Markov decision processes: a tool for sequential decision making under uncertainty, *Medical Decision Making, 30*, 474-483.

Bennett, C. C., & Hauser, K. (2013) Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach, *Artificial Intelligence in Medicine, 5*, 9-19.

Chen, X., Bailly, G., Brumby, D. P., Oulasvirta, A., & Howes, A. (2015) The emergence of interactive behavior: A model of rational menu search, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, New York: ACM, 4217-4226.

Chen, X., Starke, S. D., Baber, C., & Howes, A. (2017) A cognitive model of how people make decisions through interaction with visual displays, *Proceedings of the 35th CHI Conference on Human Factors in Computing Systems*, New York: ACM, 1205-1216.

Hamilton, W. (2010) Cancer diagnosis in primary care, *British Journal of General Practice, 60*, 121-128.

Laudicella, M., Walsh, B. Burns, E. and Smith, P. (2016) Cost of care for cancer patients in England: evidence from population-based patient-level data, *British Journal Of Cancer*, *114*, 1286.

Lee, J.D. and See, K.A. (2004) Trust in Automation: designing for appropriate reliance, *Human Factors, 46,* 50-80.

Watkins, C. J., & Dayan, P. (1992) Q-learning, *Machine Learning*, *8*, 279–292.

Wickens, C.D. and Dixon, S.R. (2007) The benefits of imperfect diagnostic automation: a synthesis of the literature, *Theoretical Issues in Ergonomics Science, 8,* 201-212.