Automated Preferences Elicitation

Miroslav Kárný
Department of Adaptive Systems
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic
school@utia.cas.cz

Tatiana V. Guy
Department of Adaptive Systems
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic
guy@utia.cas.cz

Abstract

Systems supporting decision making became almost inevitable in the modern complex world. Their efficiency depends on the sophisticated interfaces enabling a user take advantage of the support while respecting the increasing on-line information and incomplete, dynamically changing user's preferences. The best decision making support is useless without the proper preference elicitation. The paper proposes a methodology supporting automatic learning of quantitative description of preferences. The proposed elicitation serves to fully probabilistic design, which is an extension of Bayesian decision making.

1 Introduction

A feasible and effective solution of preference elicitation problem decides on the efficiency of any intelligent system supporting decision making. Indeed, to recommend a participant[1] an optimal sequence of optimal decisions requires knowing some information about what the participant (affected by the recommended decision, if accepted) considers as "optimal". Extracting the information about the participant's preferences or utility is known as *preference elicitation* or utility elicitation[2]. This vital problem has been repeatedly addressed within artificial intelligence, game theory, operation research and many sophisticated approaches have been proposed [7], [8], [6], [5]. A number of approaches has arisen in connection with applied sciences like economy, social science, clinical decision making, transportation, see, for instance, [18], [9]. To ensure feasibility and practical applicability, many decision support systems have been designed under various assumptions on the structure of preferences. In particular, a broadly accepted additive independence [16] of values of individual attributes is not generally valid. In many applications the preferences of attributes are mostly dependent and the assumption above significantly worsens the elicitation results[3].

To benefit from any decision support, the preferences should be known in the form allowing their processing by an intended decision support system. Unless the participant's preferences are completely provided by the participant, they should be learned from either *past data* or *domain-specific*

[1]Participant is also known as user, decision maker, agent.
[2]The term utility generally has a bit different coverage within decision-making context.
[3]The assumption can be weakened by a introducing a conditional preferential independence, [4].

information (technological knowledge, physical laws, etc.). Eliciting the needed information itself is inherently hard task, which success depends on experience and skills of an *elicitation expert*. Preferences can be elicited *from past data* directly collected on the underlying decision-making process or from indirect data learned from a number of similar situations. Despite acquiring the probabilistic information from data is well-elaborated, learning can be hard, especially when the space of possible behaviour is larger than that past data cover. Then the initial preferences for the remaining part of the behaviour should be properly assigned.

The process of eliciting of the *domain-specific information* is difficult, time-consuming and error-prone task[4]. Domain experts provide subjective opinions, typically expressed in different and incompatible forms. The elicitation expert should elaborate these opinions into a distribution describing preferences in a consistent way. Significant difficulties emerge when competitive/complementing opinions with respect to the same collection of attributes should be merged. A proper merging their individual opinions within a high-dimensional space of possible behaviour is unfeasible. Besides domain experts having domain-specific information are often unable to provide their opinion on a part of behaviour due to either limited knowledge of the phenomenon behind or the indifference towards the possible instances of behaviour. Then, similarly to the learning preferences from past data, the optimal strategy heavily depends on the initial preferences assigned to the part of behaviour not "covered" by the domain-specific information.

Process of eliciting information itself requires significant cognitive and computational effort of the elicitation expert. Even if we neglect the cost of this effort[5], the elicitation result is always very limited by the expert's *imperfection*, i.e. his inability to devote an infinite deliberation effort to eliciting. Unlike imperfection of experts providing domain-specific information, imperfection of elicitation experts can be eliminated. This motivate the search for a feasible *automated* support of preference elicitation, that does not rely on any elicitation expert.

The *dynamic* decision making strengthes the dependence on the preference elicitation. Indeed, the participant acting within a dynamically changing environment with evolving parameters may gradually change preferences. The intended change may depend on the future behaviour. The overall task is going harder when participant interacts with other dynamic imperfect participants within a common environment.

The paper concerns a construction of probabilistic description of preferences based on the information available. Dynamic decision making under uncertainty from the perspective of an imperfect participant is considered. The participant solves DM task with respect to its environment and based on a given finite set of opinions gained from providers of domain expertise or learned from the past data or both. The set indirectly represents the preferences in a non-unique way[6]. Additionally, the participant may be still uncertain about the preferences on a non-empty subset of behaviour. To design an optimal strategy, a participant employs Fully Probabilistic Design (FPD) of DM strategies, [10, 12] whose specification relies on the notion of *an ideal closed-loop model* which is essentially a probabilistic description of the preferences. In other words, an ideal closed-loop model describes the closed-loop behaviour, when the participant's DM strategy is optimal. FPD searches for the optimal strategy by minimising the divergence of the current closed-loop description on the ideal one. Adopted FPD implies availability of probabilistic description of the environment and probabilistic description of the past closed-loop data.

Section 2 specifies assumptions under which the automated preference elicitation is proposed within the considered FPD. Section 3 describes construction of the ideal closed-loop distribution based on the information provided. The proposed solution is discussed in Section 4 followed by the concluding remarks in Section 5.

[4]It should be mentioned that *practical* solutions mostly use a laborious and unreliable process of manual tuning a number of parameters of the pre-selected utility function. Sometimes the high number of parameters makes this solution unfeasible. Then there are attempts to decrease the number of parameters to reach an acceptable feasibility level.

[5]This effort is usually very high and many sophisticated approaches aim at optimising a trade-off between elicitation cost and value of information it provides (often decision quality is considered), see for instance [3].

[6]Even, when we identify instances of behaviour that cannot be distinguished from the preferences' viewpoint.

2 Assumptions

The considered participant deals with a DM problem, where the reached decision quality is expressed in terms of a ℓ_a-tuple of attributes $a = (a_1, \ldots, a_{\ell_a}) \in \boldsymbol{a} = \prod_{i=1}^{\ell_a} \boldsymbol{a}_i$, $\ell_a < \infty$. $\prod$ denotes Cartesian product of sets $\boldsymbol{a}_i$ the respective attribute entries belong to. The occurrence of attributes depends on an optional ℓ_d-dimensional decision $d = (d_1, \ldots, d_{\ell_d}) \in \boldsymbol{d} = \prod_{j=1}^{\ell_d} \boldsymbol{d}_j$, $\ell_d < \infty$. In the considered preference elicitation problem, the following assumptions are adopted.

A1 The participant is able to specify its preferences on the respective entries of attributes $a_i \in \boldsymbol{a}_i$ such that the most preferred value of each attribute is uniquely defined. For convenience, let the best attribute value be zero.

A2 The participant has at disposal a probabilistic model $\mathcal{M}(a|d)$, which is the probability density (pd[7]) of the attributes a conditioned on decisions d. The support of $\mathcal{M}(a|d)$ is assumed to include $(\boldsymbol{a}, \boldsymbol{d})$.

A3 The participant has[8] a joint pd $\mathcal{P}(a, d)$, describing behaviour (a, d) of the closed loop formed by the acting participant and by its environment[9]. The support of $\mathcal{P}(a, d)$ is assumed to include $(\boldsymbol{a}, \boldsymbol{d})$.

A4 The participant uses fully probabilistic design (FPD), [12], of decision-making strategies. FPD considers a specification of the ideal pd $\mathcal{I}(a, d)$ assigning high values to desired pairs $(a, d) \in (\boldsymbol{a}, \boldsymbol{d})$ and small values to undesired ones. The optimal randomised strategy $\mathcal{S}^{opt}(d)$ is selected among strategy-describing pds $\mathcal{S} \in \boldsymbol{\mathcal{S}}$ as a minimiser of the Kullback-Leibler divergence (KLD, [17])

$$\mathcal{S}^{opt} \in \operatorname{Arg}\min_{\boldsymbol{\mathcal{S}}} \int_{(\boldsymbol{a},\boldsymbol{d})} \mathcal{M}(a|d)\mathcal{S}(d) \ln\left(\frac{\mathcal{M}(a|d)\mathcal{S}(d)}{\mathcal{I}(a, d)}\right) \mathrm{d}(a, d) = \operatorname{Arg}\min_{\boldsymbol{\mathcal{S}}} \mathcal{D}(\mathcal{M}\mathcal{S}||\mathcal{I}).$$

Note that the use of FPD represents no constraints as for a classical preference-quantifying utility $\mathcal{U}(a, d) : (\boldsymbol{a}, \boldsymbol{u}) \to [-\infty, \infty)$ it suffices to consider the ideal pd of the form

$$\mathcal{I}(a, d) = \frac{\mathcal{M}(a|d) \exp(\mathcal{U}(a, d)/\lambda)}{\int_{(\boldsymbol{a},\boldsymbol{d})} \mathcal{M}(a|d) \exp(\mathcal{U}(a, d)/\lambda)\, \mathrm{d}(a, d)},\ \lambda > 0.$$

Then, the FPD with such an ideal pd and $\lambda \to 0$ arbitrarily well approximates the standard Bayesian maximisation of the expected utility [15].

3 Preference Elicitation

Under the assumptions **A1 – A4**, the addressed elicitation problem reduces to a justified, algorithmic (elicitation-expert independent) construction of the ideal pd $\mathcal{I}(a, d)$.

The following steps constitute the considered construction of the preference-expressing ideal.

S1 Each ideal pd $\mathcal{I}(a, d)$ determines marginal pds $\mathcal{I}_i(a_i)$ on the respective attribute entries $a_i \in \boldsymbol{a_i}$, $i = 1, \ldots, \ell_a$. The marginal ideal pd $\mathcal{I}_i(a_i)$ respects the highest preference for $a_i = 0$ if

$$\mathcal{I}_i(a_i = 0) \geq \mathcal{I}_i(a_i),\ \forall a_i \in \boldsymbol{a}_i. \tag{1}$$

Thus, the ideal pds meeting (1) for $i = 1, \ldots, \ell_a$ respect the participant's preferences.

S2 A realistic ideal pds (meeting (1)) should admit a complete fulfilment of preferences with respect to any individual attribute entry a_i whenever the design focuses solely on it. It is reasonable to restrict ourselves to such ideal pds as the ideal pd, which cannot be reached at least with respect to individual attributes is unrealistic.

[7]pd, Radon-Nikodým derivative [21] of the corresponding probabilistic measure with respect to a dominating, decision-strategy independent, measure denoted d.

[8]or can learn it

[9]The closed-loop model $\mathcal{P}(a, d)$ can alternatively describe a usual behaviour of other participants in similar DM tasks.

The complete fulfilment of preferences requires an existence of decision strategy $\mathcal{S}_i(d)$ such that the closed-loop model $\mathcal{M}(a|d)\mathcal{S}_i(d)$ has the marginal pd on a_i equal to the corresponding marginal $\mathcal{I}_i(a_i)$ of the considered ideal pd $\mathcal{I}(a, d)$.

FPD methodology is used to specify *realistic marginal pds*, $\mathcal{I}_i^r(a_i)$, $i = 1, \ldots, \ell_a$.

S3 The set of ideal pds $\mathcal{I}(a, d)$ having given realistic marginal pds $\mathcal{I}_i^r(a_i)$, $i = 1, \ldots, \ell_a$ is non-empty as it contains the ideal pd independently combining the expressed marginal preferences $\mathcal{I}(a, d) = \prod_{i=1}^{\ell_a} \mathcal{I}_i^r(a_i)$. Generally, the discussed set contains many pds. Without a specific additional information, the chosen pd should at least partially reflect behaviour occurred in the past. Then the adequate representant of this set is the minimiser of the KLD [22] of $\mathcal{I}(a, d)$ on the joint pd $\mathcal{P}(a, d)$. According to **A3** $\mathcal{P}(a, d)$ describes the past closed-loop behaviour and serves as the most uncertain (the least ambitious) ideal: in the worst case, the ideal pd qualifies the past behaviour as the best one. The minimiser over the set of ideal pds having marginal pds $\mathcal{I}_i^r(a_i)$, $i = 1, \ldots, \ell_a$, is described below and provides the final solution of the addressed elicitation problem.

The pds $\mathcal{I}_i^r(a_i)$, discussed in Step **S2** can be obtained as follows. Let us consider the ith entry a_i. Then ℓ_a-tuple a can be split $a = (a_{-i}, a_i)$, where a_{-i} contains all attributes except a_i and the ideal pd factorises [20]

$$\mathcal{I}(a, d) = \mathcal{I}_i(a_{-i}, d|a_i)\mathcal{I}_i(a_i). \tag{2}$$

When solely caring about the ith attribute, any distribution of (a_{-i}, d) can be accepted as the ideal one. This specifies the first factor of the ideal pd (2) as ([11])

$$\mathcal{I}_i^l(a_{-i}, d|a_i) = \frac{\mathcal{M}(a|d)\mathcal{S}(d)}{\int_{(\boldsymbol{a}_{-i}, \boldsymbol{d})} \mathcal{M}(a|d)\mathcal{S}(d)\, \mathrm{d}(a_{-i}, d)}. \tag{3}$$

This choice, complemented by an arbitrary choice of $\mathcal{I}_i(a_i)$ specifies an ideal pd on $(\boldsymbol{a}, \boldsymbol{d})$ and the strategy ${}^i\mathcal{S}(d)$ minimising KLD of the closed-loop model $\mathcal{MS}$ on it cares about the ith attribute only. For the inspected ideal pd, the optimised KLD optimised with respect a strategy $\mathcal{S}$ reads

$$\mathcal{D}(\mathcal{MS}||\mathcal{I}) = \int_{(\boldsymbol{a}, \boldsymbol{d})} \mathcal{M}(a|d)\mathcal{S}(d) \ln\left(\frac{\int_{\boldsymbol{d}} \mathcal{M}(a_i|d)S(d)\, \mathrm{d}d}{\mathcal{I}_i(a_i)}\right) \mathrm{d}(a, d). \tag{4}$$

Let us assume that there is ${}^id \in \boldsymbol{d}$ such that $\mathcal{M}(a_i = 0|\, {}^id) \geq \mathcal{M}(a_i|\, {}^id)$, $\forall a_i \in \boldsymbol{a}_i$. Then, the ideal pd $\mathcal{I}(a, d) = \mathcal{I}_i(a_{-i}, d|a_i)\mathcal{I}_i^r(a_i)$ with

$$\mathcal{I}_i^r(a_i) = \mathcal{M}(a_i|\, {}^id) \tag{5}$$

meets (1) and is the realistic marginal pd in the sense described in **S2**. Indeed, the deterministic strategy ${}^i\mathcal{S}(d) = \delta(d - {}^id) =$ pd concentrated on ${}^id$ and ideal pd $\mathcal{I}_i^l\mathcal{I}_i^r$ make the KLD (4) $\mathcal{D}(\mathcal{M}\, {}^i\mathcal{S}||\mathcal{I}_i^l\mathcal{I}_i^r)$ equal to zero, which is the absolute minimum.

The constraints (5) on the marginal ideal pds exhaust all information about the preferences available, see **A1 – A3**. It remains to select one among multitude of such ideal pds meeting (5). The minimum KLD (cross-entropy) principle [22] recommends to select the ideal pd, which minimises its KLD on a pd representing the most uncertain preference description. As discussed in **S3**, the pd describing the past history serves to this purpose. The following proposition explicitly specifies the minimiser and provides the solution of the addressed preference elicitation problem.

Proposition 1 (The recommended ideal pd) *The ideal pd $\mathcal{I}(a, d)$ describing the supplied preferences via (5) and minimising KLD $\mathcal{D}(\mathcal{I}||\mathcal{P})$, where $\mathcal{P}$ describes the past history, has the form*

$$\mathcal{I}(a, d) = \mathcal{P}(d|a) \prod_{i=1}^{\ell_a} \mathcal{I}_i^r(a_i) \tag{6}$$

$$= \frac{\mathcal{P}(d, a)}{\int_{\boldsymbol{d}} \mathcal{P}(a, d)\, \mathrm{d}d} \prod_{i=1}^{\ell_a} \mathcal{M}(a_i|\, {}^id), \;\; \textit{with } {}^id \in \boldsymbol{d} : \mathcal{M}(a_i = 0|\, {}^id) \geq \mathcal{M}(a_i|\, {}^id),\ \forall i \in \{1, \ldots, \ell_a\}.$$

Proof

The convex functional $\mathcal{D}(\mathcal{I}||\mathcal{P})$ on the convex set given by considered constraints (5) has the unique global minimum. Thus, it suffices consider weak variations of the corresponding Lagrangian functional. The pd (6) makes them equal to zero and thus it is the global minimiser searched for.

$\square$

4 Discussion

Many important features of the proposed solution (6) is implied by the fact that the constructed ideal pd reflects the relation $\mathcal{M}(a|d)$ between attributes and decisions. Specifically,

- The marginal ideal pds (5) are *not* fully concentrated on the most desirable attribute value (0), which reflects the fact that $a_i = 0$ cannot be reached with certainty.
- A specific ${}^b d$ is a bad decision comparing to other ${}^o d$ if $\mathcal{P}(a = 0|\, {}^b d) << \mathcal{P}(a = 0|\, {}^o d)$. As the closed-loop model $\mathcal{P}(a, d) = \mathcal{P}(a|d)\mathcal{P}(d)$ is a factor in (6), the decision ${}^b d$ is perceived as a bad one by the ideal pd (6) unless an unbalanced experience is faced, i.e. unless $\mathcal{P}({}^b d) >> \mathcal{P}({}^o d)$. Thus, the constructed ideal distinguishes the good and bad decisions made in past if they both occur in a balanced way.

  The danger of an unbalanced occurrence of good and bad decisions can be counteracted by modifying $\mathcal{P}(a, d)$ in order to stimulate exploration. It suffices to take it as a mixture of the closed-loop model gained from observations and of an exploration-allowing "flat" pd.
- The functional form of the ideal pd is determined by the model $\mathcal{M}(a|d)$: it is not created in an ad hoc, model independent, manner unlike utilities [16].
- It is always possible to project the constructed ideal pd into a class of feasible pds by using information criterion justified in [2, 14], if the constructed ideal pd is too complex for numerical treatment or analysis.
- The model $\mathcal{M}(a|d)$ as well as the closed-loop model of the past history $\mathcal{P}(a, d)$ can be learnt in a standard Bayesian way [1, 20]. Consequently, the preference description (6), derived from them, is learned, too.
- The involved pds can quantify the joint distributions of discrete-valued as well as continuous valued attributes. This simplifies the elicitation of preferences given jointly by categorical and numerical attributes.
- The approach can be directly extended to a dynamic DM, in which attributes and decisions evolve in time. It suffices to apply Proposition 1 to factors of involved pds.
- The construction can be formally performed even when several best (mutually ordered) attributes are admitted in a variant of Assumption **A1**. The subsequent evaluations following the same construction line are, however, harder.
- The considered preference specification is quite common but it does not cover all possibilities. For instance, an attribute $a_i \in \boldsymbol{a}_i$ may have preferences specified on a proper subset $\emptyset \neq \boldsymbol{\alpha}_i \subset \boldsymbol{a}_i$. If $a_i = 0 \in \boldsymbol{\alpha}_i$ is considered as the most desirable value of the attribute, the proposed elicitation way applies with a reduced requirement $\mathcal{M}(a_i = 0|\, {}^i d) \geq \mathcal{M}(a_i|\, {}^i d)$, $\forall a_i \in \boldsymbol{\alpha}_i$, cf. (1). Then, the proposed procedure can be used without essential changes. The real problem arises when there is no information whether the most preferred attribute is in $\prod_{i=1}^{\ell_a} \boldsymbol{\alpha}_i$ or not. Then, the participant has to provide an additional feedback by specifying a rank of the newly observed attribute with respect to the initially set values 0. The problem is tightly connected with a sequential choice of the best variant, e.g., [19].

5 Concluding Remarks

The solution proposes a methodology of automated preference elicitation of the ideal pd for a common preference specification. Covering other preference specifications is the main problems to be addressed. Also, the proposed solution is to be connected with an alternative view presented in [13], where the preference elicitation was directly treated as a learning problem and reduced to a specification of a prior pd on parameters entering environment model (and thus learnable) and parameters entering only the ideal pd (and thus learnable only via a well-specified join prior pd). The design of specific algorithmic solutions for commonly used environment models is another topic to be covered. In spite of the width of the problems hidden behind these statements, the selected methodological direction is conjectured to be adequate and practically promising.

Acknowledgments

The support of the projects GAČR 102/08/0567 and MŠMT 1M0572 is acknowledged.

References

[1] J.O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, 1985.

[2] J. M. Bernardo. Expected information as expected utility. *The Annals of Statistics*, 7(3):686–690, 1979.

[3] C. Boutilier. A POMDP formulation of preference elicitation problems. In S. Biundo, editor, *AAAI-02 Proc. of the Fifth European Conference on Planning*, pages 239–246. AAAI Press/MIT Press, Durham, UK, 2002.

[4] C. Boutilier, R.I. Brafman, C. Geib, and D. Poole. A constraint-based approach to preference elicitation and decision making. pages 19–28. Stanford, CA, 1997.

[5] U. Chajewska and D. Koller. Utilities as random variables: Density estimation and structure discovery. In *Proceedings of UAI–00*, pages 63–71. 2000.

[6] N. Cooke. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41:801–849, 1994.

[7] K. Gajos and D.Weld. Preference elicitation for interface optimization. 2005.

[8] S. Zilles H. P.Viappiani and C. Boutilier. Learning complex concepts using crowdsourcing: A Bayesian approach. In *Proceedings of the Second Conference on Algorithmic Decision Theory (ADT-11)*. Piscataway, NJ.

[9] H.B. Jimison, L.M. Fagan, R.D. Shachter, and E.H. Shortliffe. Patient-specific explanation in models of chronic disease. *AI in Medicine*, 4:191 –205, 1992.

[10] M. Kárný. Towards fully probabilistic control design. *Automatica*, 32(12):1719–1722, 1996.

[11] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař. *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, London, 2006.

[12] M. Kárný and T. V. Guy. Fully probabilistic control design. *Systems & Control Letters*, 55(4):259–265, 2006.

[13] M. Kárný and T.V. Guy. Preference elicitation in fully probabilistic design of decision strategies. In *Proc. of the 49th IEEE Conference on Decision and Control*. IEEE, 2010.

[14] M. Kárný and T.V. Guy. On support of imperfect bayesian participants. In T.V. Guy, M. Kárný, and D.H. Wolpert, editors, *Decision Making with Imperfect Decision Makers*, volume 28. Springer, Berlin, 2012. Intelligent Systems Reference Library.

[15] M. Kárný and T. Kroupa. Axiomatisation of fully probabilistic design. *Inf. Sciences*, 2011.

[16] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. JohnWiley and Sons Inc., 1976.

[17] S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–87, 1951.

[18] G. Linden, S. Hanks, and N. Lesh. Interactive assessment of user preference models: The automated travel assistant. 1997.

[19] A.A.J. Marley and J.J. Louviere. *Journal of Mathematical Psychology*, 49(6):464–480, 2005.

[20] V. Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.

[21] M.M. Rao. *Measure Theory and Integration*. John Wiley, New York, 1987.

[22] J. Shore and R. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1):26–37, 1980.