Julia Brettschneider[1]

Perspectives on decision trees

(Version 31.8.2018 (preliminary))

Abstract

We shed new light on old decision trees by introducing a mathematical framework including multiple subjective perspectives of individual decision makers simultaneously within the same decision tree model. This includes utility and knowledge that can depend on the decision maker as well as the decision point. It further includes individual awareness windows for modelling the subjective degrees of memory and foresight decision makers consider for their decisions. Both decision and games are covered by our framework.

Decision makers may act in concert, coordinated or not, or they may take turns. To model the simultaneous presence of multiple decision maker in the same decision point, we introduce the concept of influence, which is a generalisation of ownership. We propose how individual decision makers rules can be jointly implemented according to the distribution of influence. Technically, this is implemented by a probabilistic interpretation of ownership. Perspectives being depending on decision maker and decision point, this naturally leads to the concepts of utility tree, knowledge tree and awareness framework, which are structures assigning the perspective applicable in a given point according to a given influence distribution.

Decision rules are specific to the decision makers, their perspectives and the decision point. They can select the next immediate outcome only, or an outcome further down the paths, in particularly a final outcome. They can be chained together raising questions around consistency of global with local rules. Given an influence distribution, a decision rule tree is constructed assigning the relevant perspectives in a given point.

Applications of such models naturally arise in processes involving multiple people with different perspectives and evolving roles. For example, acceptance of conditional offers, medical diagnosis, individual trader investments, joint treatment decision making and population based screening.

KEYWORDS: Decision trees, games, decision makers, expected utility, joint decision making, influence, subjective probability, information.

1. Introduction

This paper develops a comprehensive and flexible language for sequential decision processes involving multiple subjective decision makers at any decision situation.

[1]University of Warwick, Department of Statistics, Coventry, UK.
email: julia.brettschneider@warwick.ac.uk

In particular, the model also includes games in the sense that decision makers take turns.

The first source of motivation for our work is the need for more detailed mathematical descriptions of the human decision making process taking into account, for example, the role of emotions and memory. For example, Gigerenzer and Goldstein **??** propose simplifications considering limitations to rationality in practical situations operating with limited resources. The review paper by Hastie gives an overview of such non-traditional developments in decision theory and sets goals for the next two decades of that area or research **??**.

A second source of motivation the need to have one framework covering processes that have the nature of sequential games as well as decisions. Seeing decisions as games against nature is a common way to unify the theory. We take the alternative approach of interpreting games as sequences of decisions with changing decision makers.

A third source of motivation is the observation that it should be able to model decision sharing should be independent of the value spaces of utility functions. This allows a very general notion of shared decision making that can be implemented independently of the value space of a utility function.

A forth source of motivation is Shafer's distinction of decision points into chance, choice and intermediate situations. Bringing this together with the unified decision-game framework leads simultaneoul

Game trees and decision tree goes back a long time (e.g. [**?**]). One of the players is nature means decision tree. It is a game tree with probabilities supplied for only for some of the branchings. This leads to the common visualisation of decision trees with two types of vertices: squares for branchings corresponding to *decision situations* and circles for *chance situations.* Shafer [**?**] also allows *intermediate situations*, which are branchings with partial probabilisation. Markov decisions processes (MDP) and their partially observed counterparts (POMDP) are framework that build on this. They are very constraint in that they require an alternation of decision situations and chance situations. In particular, Shafer's generalised situations are not included.

We consider models that are very flexible about the order and the type of the decision situations. Decision processes are represented by finite rooted trees (Section 2). Each generation of the tree corresponds to a step in the decision process. The vertices of the tree correspond to intermediate and final outcomes of the decision process. Decision makers are equipped with influence on a scale between 0 for *no influence* to 1 for full control (Section 4). Both concepts, decision ownership and decision influence, who makes decisions is not hardwired by the layout of the tree branches, but represented by functions of the tree.

In the Sections 3 and **??** we will develop a unified framework modelling characteristics of situations including ownership of a decision and knowledge about the probability of outcomes. In particular, it covers the thee three situations considered by Shafer mentioned above and can also be used to describe ambiguity. The technique is to go on a higher level making the decision maker probabilistic. In other words, a random DM is chosen governed by a particular distribution. Then, the relevant decision role will be used. In special cases where the value space has more structure, there will be be rules for how to make compromises between a group of decision makers.

In a traditional sequential game, decision makers take turns being in full control of a decision step. We want to generalise this by allowing partial control which we call influence. The decision space is fix, that is the tree can not grow new branches for compromises not in the scope of this papers. Without assuming any additional structure on V, however, Hence need to go a level up. First decide who make the decision, then that person takes it. Influence in this context means the likelihood to be chose to make the decision, which is given a probability distribution.

Section **??** explains the subjective ingredients of the decision process. This includes utility and knowledge, which are specific to the decision maker. They can be used in flexible ways, evaluated just the one relevant for the decision maker(s) in charge, as defined by the influence tree. In addition, we introduce the notion of awareness that governs who much past and future a decision maker considers. It is depending on both the decision maker and the decision point. This allows to analyse the impact of changes in memory and foresight independently from other changes in the utility and knowledge functions.

Section 6 puts this all together in one decision model and builds decision rules on it. We distinguish different degrees of locality up to final decision rules selecting final outcomes. Inspired by Smith's notion of introduces the notion of currently judgement optimal [**?**] we look at consistency between local and global decision roles. An important concept is adaptation, which refers to the use of whatever is available in the awareness range of a decision point.

2. Trees

Trees have long played a central role in probability theory as well as decision theory; a good review of their usefulness can be found in [**?**].

Mathematically, a tree is a finite connected graph without cycles and with a distinguished vertex ρ also called its *root*. Let $\mathcal{T}$ denote the set of all trees. The root describes generation 0, its neighbours form generation 1, their other neighbours form generation 2, and so on.

Let $T \in \mathcal{T}$. Any two vertices $x, y \in T$ are connected by exactly one path $\nu_T(x, y)$ and their distance $d_T(x, y)$ is defined as the number of its edges. Let $g(x)$ be the generation of the vertex x. For any $n \in \mathbb{N}_0$, let $g_n(T) = \{x \in T \,|\, g(x) = n\}$ be the nth generation of the tree.

The height of the tree is defined as the maximal possible generation and its diameter is the maximal possible distance between vertices

$$\mathrm{ht}(T) = \max_{x \in T} g(x) \qquad \mathrm{diam}(T) = \max_{x,y \in T} d_{\mathcal{T}}(x, y)$$

By definition, $g_{\mathrm{ht}(T)}(T) \neq \rho$, but not all paths starting from the root and heading directly to a leaf need to be of length $\mathrm{ht}(T)$. The set of all possible individuals of such trees is given by

$$\mathbb{T} = \rho \cup \bigcup_{n \in \mathbb{N}} \mathbb{N}^n.$$

For a vertex $y \in g_n(\mathbb{T})$ with $n \in \mathbb{N}$, there exists exactly one vertex $x \in g_{n-1}(\mathbb{T})$ that is connected to y. The vertex y is called a *child* of the *mother* x and $m(y) = x$ defines the *mother map* $m : g_n(\mathbb{T}) \longmapsto g_{n-1}(\mathbb{T})$. $C(x) = \{y \in T \,|\, m(y) = x\}$ is the set of children of x. Iterating this returns the set of *grand children* C^2, the set of *great grand children*

A vertex $x \in T$ is called *leaf* if $|C(x)| = 0$, and the set of leafs is called $L(T)$. Note that not all paths starting from the root and heading directly to a leaf need to be of length $\mathrm{ht}(T)$, but some can be shorter. Let $T' = T \setminus L(T) = \{x \in T \,|\, |C(x)| > 0\}$ be the tree without its leaves. The mother map can be iterated to identify x's grandmother $m^2(x)$, great-grandmother $m^3(x)$ and so on.

The mother map defines a relation on $\mathbb{T}$ by

$$x \preceq y \quad \Leftrightarrow \quad m^n(y) = x \text{ for some } n \in \mathbb{N}_0$$

Since $\preceq$ is reflexive, antisymmetric and transitive, $(\mathbb{T}, \preceq)$ is a partially ordered set. ρ is the least element. The individuals of any particular tree $T \in \mathcal{T}$ form a finite subset of $\mathbb{T}$, more specifically,

$$T \subseteq \; \rho \; \cup \bigcup_{\substack{n \in \mathbb{N} \\ n \leq \mathrm{ht}(T)}} \mathbb{N}^n = \; \rho \cup g_1(\mathbb{T}) \cup g_2(\mathbb{T}) \cup \cdots \cup g_{\mathrm{ht}(T)}(\mathbb{T}).$$

Since the mother map can be restricted on any $T \in \mathcal{T}$, $x \preceq y$ is inherited and defines a partial order on T. This coincides with a traditional definition of a partial order on a tree that is defined by a circuit free path from ρ to y going through x.

For each fixed $y \in T$, the set $\{x \in T \mid x \preceq y\}$ is totally ordered by $\preceq$ and we define

$$d_y(x) = |\{x \in T \mid x \preceq y\}|. \tag{1}$$

For $n \in \mathbb{N}$, the projection map

$$\pi_n : \mathbb{T} \longrightarrow \bigcup_{\substack{k \in \mathbb{N}_0 \\ j \leq n}} g_j(\mathbb{T}) \qquad \text{with} \qquad \pi_n(x) = \begin{cases} m^{g(x)-n}(x) & \text{if} \in g(x) - n \geq 0, \\ x & \text{otherwise.} \end{cases} \tag{2}$$

assigns vertices belonging to generations later than n their ancestor in generation n. This can be applied to whole tree $T \in \mathcal{T}$, resulting in a tree $\pi_n(T)$ of height n. In the opposite direction, we define the subtree tree with root x and continuing all the way to the leaves, i.e., the tree containing x and all its descendants

$$\hat{\pi}(T)(x) = \{y \in T \mid y \succeq x\}. \tag{3}$$

We also define smaller subtrees through restricting it to only n post-x generations

$$\hat{\pi}_n(T)(x) = \{y \in \hat{\pi}(T)(x) \mid d_y(x) \leq n\} = \{y \in \hat{\pi}(T)(x) \mid g(y) \leq g(x) + n\}. \tag{4}$$

The *maximal generation of the descendants* of x is given by

$$\max\{g(y) \mid y \in \hat{\pi}(T)(x)\}. \tag{5}$$

A tree is $T' \in \mathcal{T}$ called *subtree* of T if $T' \subseteq T$ and *supertree* of T if $T' \supseteq T$.

We need systematic notation to refer to the individual vertices. The classical Ulam-Harris framework (see e.g. Jagers) identifies vertices with sequences of natural numbers so that $x = (x_1, x_2, \ldots, x_n) \in \mathbb{N}^n$ denotes the x_nth child of the x_{n-1}th child of the ... and so on up to the x_2th child of the x_1th child of the root ρ.

A tree can be used to describe a structure of events, also called *event tree*. Equipping the branches with probabilities is referred to as *probabilising* the tree, making it a *probability tree* (see e.g. Edwards 83).

Trees can be used to describe (sequential) games and decision processes, with the tree's generations representing steps of the process. The leaves $L(T)$ correspond to *final outcomes*. T' contains all other vertices which double function as *decision points* and *intermediate outcomes*. Traditionally, the latter are also called *situations*. For each decision point $x \in T'$, the choices are given by the set $C(x)$ of its children. A function $r : T \longrightarrow \mathcal{R}$ assigns the outcome of each decision point a *raw value* from a set $\mathcal{R}$.

3. Ownership

Decision processes are controlled decision makers. We will define mathematical concepts for different levels of controls. In this section we consider cases where each decision point is owned by exactly one decision maker.

Let B be a set of decision makers. This is normally finite, but it may be countable in some applications. A decision maker would typically be a person, but the role could also be taken on by an institution, a robot or an external force (of which little may be known). The latter is often referred to as *nature* in the economics literature.

Definition 1. Decision owner and control tree. *Let $T \in \mathcal{T}$ be a decision tree and $\beta : T' \longrightarrow B$ a map assigning each decision point x a decision owner $b(x)$. $\beta^T := \beta(T')$ is called* ownership tree *generated by β.*

Note that it is justified to refer to β^T as a tree because it inherits the structure of T'.

In the simplest case one decision maker owns every decision, which has the side effect of eliminating uncertainty from the decision process.

Example 2. Full control. *If there is a decision maker $b \in B$ such that $\beta(x) = b$ for all $x \in T'$ then b fully owns or has full control over the decision process.*

The emphasis of this paper are decision processes that involve more than one decision owner, including non-human ones, and that allow them to take turns or share ownership of decisions. This includes classical set ups such as a sole decision owner making decisions in the presence of uncertainty, or games with several players.

An important subclass is when control over entire steps of the process is assined to fixed decision makers. Note that this is not generally the case, because who controls the next step could depend on the outcomes of previous steps. For example, someone who made a decision leading to a positive outcome one or more steps later, may gain influence (see Example ???).

Definition 3. Control of a step. *Let $n \in \mathbb{N}_0$. If there is a $b \in B$ such that $\beta(x) = b$ for all $x \in g_n(T')$, than b* controls the nth step *of the decision process.*

Definition 4. Sequential control. *If for any $n = 0, \ldots, ht(T')$ there is a decision maker $b \in B$ such that b controls the nth step then the decision process is* sequentially controlled.

Example 5. Traditional 2-person sequential game. *N_1 are be odd numbers and N_2 are the even numbers and Let $B = \{b_1, b_2\}$ and for $x \in T'$ let*

$$\beta(x) = \begin{cases} b_1 & \textit{if } x \textit{ is odd,} \\ b_2 & \textit{if } x \textit{ is even.} \end{cases}$$

For example, a sequential game of with two rounds under each players control is represented by a suitable tree T of height 4 with $N_1 = \{1, 3\}$ and $N_1 = \{0, 1\}$.

A more general form of sequential games is when there are finitely many players, one of them owning the decision in each step, though not necessarily following a fixed repeated pattern. This corresponds to the definition of a sequential decision process 4.

Example 6. Generalised sequential game. *Let* $N_j \subset \mathbb{N}_0, j \in \{1, \ldots, J\}$, *be a partition of* $\mathbb{N}_0$, $B = \{b_1, \ldots, b_J\}$ *and for* $x \in T'$ *let* $\beta(x) = b_j$ *if* $x \in g_{N_j}$.

A property we want to highlight specifically is if two decision owners complement each other. This can be aligned with steps, as is the case in a sequential game examples, but it can also defined for more general situations, where decision ownership is more dependent on outcomes, as the example following the definition.

Definition 7. Complementary ownership. *Let* B *be a set of decision makers and* $b_1, b_2 \in B$. *Two decision makers* b_1, b_2 *own complementary parts of a decision process if*

$$\beta(x) = b_1 \qquad \Longleftrightarrow \qquad \beta(x) \neq b_2 \qquad \text{for all } x \in T'.$$

Example 8. Decision sharing. *This is a 3-step decision process where two people* b_1 *and* b_2 *share decision ownership (e.g. two co-workers dividing up tasks by competence), while a third one (e.g. an external force)* b_3 *also controls some of the decisions.*

$$\begin{aligned} \beta(\rho) &= b_1, \\ \beta(1) = b_2, \beta(2) &= b_3, \\ \beta(11) = b_3, \beta(12) = b_1, \beta(21) = b_2, \beta(22) &= b_1. \end{aligned}$$

In some models, one or more of the decision makers are distinguished as representing a form of external control. For example, *games against nature* in economic game theory fall into this category. This and more general layouts are covered by the above formalism, with one of more coordinates of β representing one or more of such forces. For illustration, here is a concrete example of a 3-step sequential decision process that provides a template that involves two decision makers and one external force.

Example 9. *Let* $T \in \mathcal{T}$ *with* $ht(T) = 3$. *Let* b_1 *and* b_2 *be human decision makers and* b_3 *be an external force. In the first step of this decision process,* β_2 *decides whether or not to make a conditional offer to* β_1. *In the second step of the decision process,* β_1 *decides whether or not to accept it. In the third step, an external force decides whether or not the condition of the offer is fulfilled. The control tree is* *given by*

$$\beta(x) = \begin{cases} b_1 & \textit{for all } x \in \rho, \\ b_2 & \textit{for all } x \in g_1(T'), \\ b_3 & \textit{for all } x \in g_2(T'). \end{cases}$$

All pairs of decision makers in this example are complementary to each other.

Example 10. Examples for non-sequential process, ideally something relevant for applications later. For example, a class of examples representing investigations of some sort. Pick up again in last section, e.g. criminal case analysis (maybe use concrete example from Sherlock Holmes in last section), medical diagnostics, environmental pollution, loss of water in a distribution system, repairing broken machine. Could later have decision maker dependent on external force (under bad conditions) or in one of the decision markers control, could also depend on covariates (e.g. A-level results based on private vs state).

So far we have described a crude way of assigning control over decisions in absolute terms, full or none. For many situations, this is too rigid and the next section is devoted to introducing a softer notion of decision ownership that we will call *decision influence.*

4. Influence

To start with, we define influence of decision makers on a decision at point $x \in T'$. Without further structural properties on the set of outcomes $r(C(x)) \subseteq \mathcal{R}$ it is not straight forward how to do that. In cases where $\mathcal{R}$ is ordered or equipped with a metric (e.g. $\mathbb{R}$), influence could be expressed through combinations of different decision makers' preferences. Such quantitative influence notions will be discussed further below. For a general spaces $\mathcal{R}$, the key idea is to add an additional layer to the decision process. The latter will govern the selection of a decision maker in decision point $x \in T'$ via a probability distribution who will control that decision.

Definition 11. Influence distribution and influence tree. *Let $T \in \mathcal{T}$ be a decision tree and $P = (p_x)_{x \in T'}$ be a family of probability distributions on a set of decision makers B. For each $x \in T'$ let β_{p_x} be a random variable with distribution p_x. Then P is called* influence distribution *and β_P defined by $\beta_P(x) = \beta_{p_x}$ $(x \in T')$ is called* influence tree *generated by P.*

Note that in this set up, the influence decision makers have in a particular decision point is independent of their influence on other decision points, but that could be generalised by replacing the pointwise construction of β_P through marginals by a construction that allows dependencies.

The definition of influence as a quantitative concept is justified by an interpretation as average. Any given realisation of the random variable β_P defines an ownership tree as in Section 3 with only one decision maker per decision point. Let $(\beta^{(i)})_{i\in\mathbb{N}}$ be a sequence of independent realisations of the same influence tree β_P generated by P. Then, by Borel's law of large numbers, with probability 1,

$$\frac{1}{n}\Big|\big\{i \in \mathbb{N} \,\big|\, 1 \leq i \leq n, \beta^{(i)}(x) = b\big\}\Big| \longrightarrow p_x(b) \quad \text{for } n \to \infty$$

for all $b \in B$ and for each $x \in T'$. That means, the probability that the decision in x is taken by decision maker b is $p_x(b)$, which justifies the interpretation that b has an influence of $p_x(b)$ on decision point x.

The concepts ownership and influence can formally be linked by representing ownership as a special case of influence using point measures.

Example 12. *[Given an ownership tree β define the $p_x = \delta_{\beta(x)}$ $(x \in T')$. Then the corresponding influence tree reduces to the original ownership tree:*

$$\beta_P(x) = \beta_{\delta_{\beta(x)}} = \beta(x) \ (x \in T').$$

Some of the theory and examples below will just be spelled out for influence as the case of ownership can be derived from this.

In analogy to Definition 4 consider the special class of sequential decision processes. Probability distributions governing the selection of the decision makers simply need to be applied to entire steps of the process rather than to the individual vertices.

Definition 13. Sequential influence. *Let $T \in \mathcal{T}$ be a decision tree and $P = (p_n)_{n=0,\ldots,ht(T')}$ be a family of probability distribution on B. For $n = 0, \ldots, ht(T')$ let β_n be a random variable with distribution p_n. Then P is called sequential influence distribution and β_P defined by*

$$\beta_P(x) = \beta_n \textit{ for all } x \in g_n(T') \quad (n = 0, \ldots, ht(T'))$$

is called sequential influence function generated by P.

5. Perspectives

Different decision makers see the process from different perspectives and have different windows of attention. This is being formalised by including multiple realisations of traditional decision theory concepts such as utility and knowledge in the same decision model and by introducing awareness to allow for more flexible modelling of the individual's selected attention.

5.1. **Utility.** It is common to replace the raw value of a decision outcome by an interpretation involving a utility function u. One important modification is to ensure that the appropriate subjective interpretation is used, that is, the utility used by the owner of the decision point, either in the deterministic sense (i.e., based on an ownership tree) or in a probabilistic sense (i.e.,based on an influence tree). Traditionally, the utility only depends on the final outcome, but we define utility also for intermediate outcomes of the decision process, which can be achieved by defining a utility $u : T \longrightarrow \mathbb{R}$ for the whole tree. In a prospect theory paradigm, the utility function is replaced by the value function. The concepts developed here can be applied to this situation as well.

In empirical examples, all ingredients discussed here may be estimated from data. This can include covariates that may depend on the situation and the decision maker.

Definition 14. Utility tree. *Let $T \in \mathcal{T}$ be a decision tree equipped with an influence tree β_P generated by P. For each $b \in B$ let $u_b : \mathcal{R} \longrightarrow \mathbb{R}$ be a the utility of decision maker b. Then the* utility tree *U is defined as*

$$U(x) = u_{\beta_{px}}(r(x)) \qquad (x \in T).$$

It is justified to refer to U as a tree, because the structure is inherited. The utility in $x = \rho$ can also be thought of as a reference or initial state of the utility. In all decision points the utility can depend on the decision maker. This allows to consider decision makers with different types of utilities a parts of the same decision model.

The definition is formulated for the general case of an influence tree. For the case of an ownership tree the above construction simplifies as follows.

Example 15. Utility tree in the ownership case. *Let $T \in \mathcal{T}$ be a decision tree equipped with an ownership tree β on B. For each $b \in B$ let $u_b : \mathcal{R} \longrightarrow \mathbb{R}$ be a the utility of decision maker b. Using the correspondence in Example 12 the Utility tree U reduces to*

$$U(x) = u_{\beta(x)}(r(x)) \qquad (x \in T).$$

In many traditional examples of decision processes, only the final outcomes matter. The next example capture this case. We keep it specific to decision makers, which allows to have both decision makers consider intermediate outcomes utilities and those who only consider final outcomes as part of the same model.

Example 16. Final outcomes utility. *A decision maker b who only takes into account the utility of the final outcome is modelled by a utility function of the form*

$$u_b : T \times \mathbb{R} \longmapsto \mathbb{R} \text{ with } u(x, r) = 0 \text{ for all } x \in T \setminus L(T).$$

The following class of examples captures the opposite situation.

Example 17. Elephant utility. *Assume the utility tree U has the property*

$$U(x) = \sum_{\rho \preceq y \prec x} U(y) \text{ for all } y \in T.$$

This describes a utility that is build up by summing up all utility accumulated along the way. Concrete examples for this can easily be constructed iteratively.

Real world situation may be somewhere between these two extremes. Decision rules can be restricted to considering a certain number of past steps only.

In Definition 14 utility trees were based only on the raw values using utility functions that live there, but that way we can not track which decision situation is being considered. The concept of a utility tree can also be generalised further to allow dependency on the decision situation itself.

Definition 18. Situation dependent utility tree. *Let $T \in \mathcal{T}$ be a decision tree equipped with an influence tree β_P generated by P. For each $b \in B$ let $u^\star_b : T \times \mathcal{R} \longrightarrow \mathbb{R}$ be the situation dependent utility of decision maker b. Then the* situation dependent utility tree *U^* is defined as*

$$U^\star(x) = u^\star_{\beta_{p_x}}(x, r(x))) \qquad (x \in T).$$

Definition 14 is a special case of Definition 18 using $u^\star_b(x, \cdot) = u_b$ for all $x \in T$.

Apart from allowing preferences depending on decision makers and situation, rather than raw values only, utility also opens up more options to define compromises between different decision makers simultaneously influencing the same decision situation. As discussed at the beginning of Section 4, without making additional assumptions on $\mathcal{R}$, such combinations would be done via a probabilistic combination of their individual decision. Having carried over their individual preferences to $\mathbb{R}$ via utility functions, the can not be combined by rules involving arithmetic operations. For example, a rule for decision in point x could be to choose the child of x that maximises the utility over all children or x and over all decision makers. (This assumes sensibly normalised utility across decision makers.) More examples will be discussed in detail in the next section.

5.2. **Knowledge.** The second aspect that distinguishes decision makers from each other is the knowledge available to them. We assume knowledge is described by values from a set $\mathcal{S}$ and can be discovered while moving through the decision process. We present knowledge using the same tree based notation as for the utility, allowing dependency on the decision situation and the decision maker. In practical applications, knowledge can e.g. be used to include covariates of an observed process. Per definition, knowledge would naturally include the path taken in the decision process so far, though one could choose to ignore this. We use the knowledge rather than information, because the latter is often used in decision theory something particular gained through a specific experiment and quantified through the impact it has on decision making in the sense of deGroot rather than in Lindsey's interpretation of a general added value (see e.g. [**?**]). The knowledge concept allows both interpretations.

Definition 19. Knowledge tree. *Let $T \in \mathcal{T}$ be a decision tree equipped with an influence tree β_P generated by P. For each $b \in B$ let $\kappa_b : T \longrightarrow \mathcal{S}$ be a function that assigns each $x \in T$ the knowledge available to decision maker b in that decision point. Then the* knowledge tree *K is defined as*

$$K(x) = \kappa_{\beta_{p_x}}(x) \qquad (x \in T).$$

5.3. **Awareness.** The last ingredient distinguishing different decision makers' perspectives is their awareness of the tree and properties attached to it. For reasons including both cognitive capacity and emotional constraints they may not use the full tree for their decision making. We model the awareness separately rather than hardwiring it into the tree itself to have the flexibility to study the effects of knowledge selection on the decision outcomes.

This includes their degree of foresight as well as their depth of memory. Both can be path dependent rather than of a fixed level for each decision point.

Definition 20. Memory. *A function $\psi^- : T \longrightarrow \mathcal{P}(T)$ on a decision tree $T \in \mathcal{T}$ is called* memory function *if for any $x \in T$, $\psi^-(x)$ is connected and $x \in \psi^-(x)$.*

Definition 21. Foresight *A function $\psi^+ : T \longrightarrow \mathcal{T}$ on a decision tree $T \in \mathcal{T}$ is called* foresight function *if for any $x \in T$, x is the root of $\psi^+(x)$ and $\psi^+(x) \subseteq T$.*

This can be combined to for all decision points and decision makers involved.

Definition 22. Awareness. *Let $T \in \mathcal{T}$ be a decision tree equipped with an influence tree β_P generated by P. For each $b \in B$ let ψ_b^- the memory function of band ψ_b^+ the* memory and foresight trees *are defined as*

$$\Psi^-(x) = \psi^-_{\beta_{p_x}}(x) \text{ and } \Psi^+(x) = \psi^-_{\beta_{p_x}}(x) \text{ for } (x \in T).$$

The combination $\Psi = (\psi^-, \psi^+)$ is called awareness range.

To look at some examples, we start with decision makers who consider all of the past or all of the available future options.

Example 23. Maximal memory. *A memory function ψ^- with*

$$|\psi^-(x)| = g(x) \text{ for all } x \in T$$

describes maximal memory, *that is, a decision maker who remembers everything from the beginning of the decision process.*

Example 24. Maximal foresight. *A foresight function* ψ^+ *with*

$$L(\psi^+(x)) \subseteq L(T) \text{ for all } x \in T$$

describes maximal foresight, *that is, a decision maker who looks as many steps ahead as possible.*

The amount of memory or foresight can be limited to a fixed number of steps.

Example 25. k-step memory. *Let* $k \in \mathbb{N}_0$.

$$\psi^-(x) = \{y \in T \mid y \preceq x, d_x(y) \leq k\} \text{ for all } x \in T$$

defines a memory function that describes a decision maker who remembers the last k *steps of the decision process (or as many as are available).*

Example 26. k-step foresight. *Let* $k \in \mathbb{N}_0$.

$$\psi^+(x) = \{y \in T \mid y \succeq x, d_y(x) \leq k\} \text{ for all } x \in T$$

defines a foresight function that describes a decision maker who looks k *steps (or as many as are available).*

This includes decision makers who quickly forget or look at most one step ahead, as well as the extreme cases of ignoring all past or future knowledge.

Example 27. Forgetful. *A* 1*-step memory function describes a decision maker who only remembers one past decision situation.*
Amnesia. *A* 0*-step memory function describes a decision maker who forgot the past entirely and can only rely on the context given about the present decision point (and potentially also context of the future).*
Myopic. *A* 1*-step foresight function describes a decision maker who only looks (up to) one step ahead.*
No future. *A* 0*-step foresight function describes a decision maker who considers no future context whatsoever.*

6. Decisions

We now introduce decision rules incorporating multiple decision makers with individual perspectives. This is build onto the concepts introduced in the last section. To summarise, our basis for the remaining paper is:

Model for a decision process with perspectives: A decision process is described by the collection $(T, B, \beta_P, U, K, \Psi)$ with

- T a tree including all decision points,
- B a set of decision makers,
- β_P an influence tree generated by an influence distribution P on B,
- U a utility tree,
- K a knowledge tree,
- Ψ an awareness framework.

If there is only one decision maker b this takes the form $(T, \{b\}, Id_{\delta_b}, u_b, \kappa_b, \Psi_b)$ and we write for short (T, b, u, κ, Ψ).

The utility U and the knowledge K contain the perspectives decision rules are based on. As they are specific to the decision maker in charge, β_P governs which perspectives are used at any given position and Ψ selects which selection of it.

Decision rules can be deterministic or (fully or partly) stochastic making their outcomes random variables. They can be local in the sense of selecting only the next step of the decision making process, or they can cover several or all of remaining steps.

First, we look at just one decision maker.

Definition 28. Decision rule for an individual decision maker. *Consider a single decision maker model* (T, b, u, κ, Ψ)*. A* k-step decision rule $D_x(u, \kappa)$ *in* $x \in T'$ *assigns the decision point* $x \in T'$ *an outcome in* $C^k(x)$ *depending on* u *and* κ*. A* 1-step decision rule *is also called* single step decision rule. $D_x(u, \kappa)$ *is called* adapted *in* x *if*

$$D_x(u, \kappa) = D_x(u(\Psi(x)), \kappa(\Psi(x))).$$

$D(u, \kappa) = (D_x(u, \kappa))_{x \in T}$ *is called* adapted *if* $D_x(u, \kappa)$ *is adapted for all* $x \in T$. $D_x(u, \kappa)$ *is called* final decision rule tree *if its outcomes are leaves.*

Note that being adapted does not mean that the decision can only depend on the utility and knowledge in the decision point x itself, but does allow dependencies on the whole awareness range $\Psi(x)$. In other words, being adapted means that it only depends on those parts of the utility and knowledge tree that were selected by the memory and foresight functions in that decision point.

Decision rules can be chained together, e.g. for two of them $D_{D_x(u,\kappa)}(u, \kappa)$. However, this does not necessarily give the same result as $D_x^2(u, \kappa)$, because intermediate results may exclude some of the originally available paths. Adaptation does not change this. In fact, the updating of $u(\Psi(x))$ and $\kappa(\Psi(x))$ by moving a step ahead can make even more of a difference.

Here are some examples for single-step decision rules.

Example 29. Simple single-step decision rules.
Assume Ψ has foresight of at least 1 step, i.e. $\psi^+(x) \geq 1$ for all $x \in T$.
Maximum utility: $D_x(u, \kappa) = \text{argmax}\{u(y) \mid y = C(x)\}$

The traditional minimax and maximax rules are two-step decision roles obtained by chaining together two single-step decision rules.

Example 30. Two-step decision rules.
Assume Ψ has foresight of at least 2 steps, i.e. $\psi^+(x) \geq 2$ for all $x \in T$.
Maximum expected utility: $D_x(u, \kappa) = \text{argmax}\{E[u(z)] \mid z = C^2(x)\}$

Example 31. Decision stress.
Assume Ψ has full memory, i.e. $\psi^-(x) = g(x)$ for all $x \in T$, and foresight of at least 2 steps. Let κ be a function that records the mental effort the decision maker is investing and let $\alpha > 0$ be some threshold. For example, a simple surrogate value for the amount of decision stress is the number of options considered so far at each step:

$$\kappa(x) = \sum_{n=1}^{g(x)-1} |C(x_1, x_1, \ldots, x_n)|.$$

This allow to model a decision maker who would normally use a rational approach, e.g. expected utility maximisation, but once exhausted just selects a random option from the available choices.
Exhausted decision maker:

$$D_x(u, \kappa) = \begin{cases} \text{argmax}\{E[u(z)] \mid z = C^2(x)\} \text{ for} & \kappa(x) \leq \alpha \\ \text{Unif}(C(x)) & \kappa(x) > \alpha \end{cases}$$

Finally, we generalise Definition 28 to the model with multiple decision makers.

Definition 32. Decision rule for multiple decision makers. *Consider a multiple decision maker model* $(T, B, \beta_P, U, K, \Psi)$. *A* k-step decision rule $D_x(U, K)$ *in* $x \in T'$ *assigns the decision point* $x \in T'$ *an outcome in* $C^k(x)$ *depending on* U *and* K. *A* 1-step decision rule *is also called* single step decision rule. $D_x(U, K)$ *is called* adapted *in* x *if*

$$D_x(U, K) = D_x(U(\Psi(x)), K(\Psi(x))).$$

$D(U, K) = (D_x(U, K))_{x \in T}$ *is called* adapted *if* $D_x(U, K)$ *is adapted for all* $x \in T$. $D_x(U, K)$ *is called* final decision rule tree *if its outcomes are leaves.*