Decision making in stochastic extensive form I: Stochastic decision forests

E. Emanuel Rapsch[a]

[a] *Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany*

18 April 2024

Abstract

A general theory of stochastic decision forests reconciling two concepts of information flow – decision trees and refined partitions on the one hand, filtrations from probability theory on the other – is constructed. The traditional "nature" agent is replaced with a one-shot lottery draw that determines a tree of a given decision forest, while each "personal" agent is equipped with an oracle providing updates on the draw's result and makes partition refining choices adapted to this information. This theory overcomes the incapacity of existing approaches to extensive form theory to capture continuous time stochastic processes like Brownian motion as outcomes of "nature" decision making in particular. Moreover, a class of stochastic decision forests based on paths of action indexed by time is constructed, covering a large fraction of models from the literature and constituting a first step towards an approximation theory for stochastic differential games in extensive form.

Keywords: Extensive form games, Dynamic games, Stochastic games, Decision making, Sequential decision theory, Stochastic processes

JEL: C73, D81

2020 MSC: 91A15, 91A18, 91B06

Contents

Introduction **2**

1 Decision forests **5**
- 1.1 Order- and graph-theoretic conventions . . . 5
- 1.2 Decision forests . . . 6
- 1.3 Forests of decision trees . . . 8

2 Stochastic decision forests **9**
- 2.1 Exogenous scenario spaces . . . 9
- 2.2 Stochastic decision forests . . . 10
- 2.3 Simple examples . . . 12
- 2.4 Action path stochastic decision forests . . . 13

3 Exogenous information **17**
- 3.1 Filtrations in probability theory . . . 17

Email address: `rapsch@math.tu-berlin.de` (E. Emanuel Rapsch)

3.2 Exogenous information structures . . . 18
3.3 Simple examples . . . 19
3.4 Action path stochastic decision forests . . . 20

4 Adapted choices **21**
4.1 Choice under uncertainty and adapted processes . . . 21
4.2 Choices in stochastic decision forests . . . 22
4.3 Simple examples . . . 23
4.4 Action path stochastic decision forests . . . 25

Conclusion **28**

Acknowledgements **29**

References **29**

Appendix A Proofs **32**
Appendix A.1 Section 1 . . . 32
Appendix A.2 Section 2 . . . 34
Appendix A.3 Section 3 . . . 40
Appendix A.4 Section 4 . . . 41

Introduction

This is the first of a series of papers whose objective is to provide a unified theory of stochastic games and decision problems in extensive form. Representing a decision problem in extensive form means specifying it in terms of what I call "extensive form characteristics", namely, the flow of information about past choices and exogenous events, as well as the set of local choices available to decision makers. Although the classical theory developed by von Neumann and Morgenstern in [45] and Kuhn in [30, 31] is bound to strong finiteness assumptions, the concept itself is very general and encompasses a large amount of applications: in a series of papers among which are [3, 4, 2] and which the monograph [5] builds upon, Alós-Ferrer and Ritzberger develop an abstract and utterly general theory of extensive form games and decision problems and give a concrete order-theoretic characterisation of its own boundaries. It follows that beyond these boundaries, the notions of strategy, randomisation, outcome, and equilibrium – if used in respect to extensive form characteristics of the decision problem at hand – have no rigorous decision-theoretic meaning.

Hence, any decision problem and game exhibiting extensive form characteristics lies within these boundaries, or is at least some sort of limit of objects within these boundaries. The inclusion, or at least the precise meaning of this limit, permits to rigorously state the decision-theoretic meaning of pure and behaviour strategies, outcomes, equilibrium etc. for the given decision problem. This opens a new perspective on problems in continuous time in particular, a sphere in which the subtlety of this issue has been known for a long time (see, e.g., [42, 44] and the references therein). Based on the maximality result in [44], systematically underlined in [4, 2], that urges to restrict outcomes to certain piece-wise constant paths, in [1] a rigorous extensive form foundation of continuous time games and decision problems with such outcomes is proposed for the first time.

How can a stochastic game exhibiting "extensive form characteristics" be modelled? The standard reply from game and decision theory is: introduce a "nature" agent executing a behaviour strategy (see, e.g., [3, section 2.2.3], or [21, 5], going back to Shapley's work [39]). But here is a fundamental problem in the continuous time case since many relevant modelling applications, from engineering (see [15, 26] and the references therein) over economics (see, e.g., [35, 24]) and finance (see [17, 34, 9, 32]) to reinforcement learning (see, for example, [25]), demand exogenous noise given by stochastic processes whose paths are all but piece-wise constant, the paradigmatic case being Brownian motion (see, for instance, [34, 29, 14] for an overview over both theory and applications). By no means this assumes the "world to be Brownian" or the like; however, such models are used extensively, be it because it appears plausible to permit unpredictably small time lags until the next exogenous information revelation, be it because certain quantities arguably evolve in decision paths of low regularity, be it for mathematical and computational convenience. This creates the fundamental issue of a lack of a proper decision-theoretic foundation of a large class of relevant models. Nevertheless, one may certainly argue that these models exhibit "extensive form characteristics" without being strictly captured within the Alós-Ferrer–Ritzberger framework ([5]).

Despite this, stochastic control and differential game theory based on stochastic analysis have found a pragmatic way to deal with a continuous time decision problem's "extensive form characteristics" (see, e.g., [14, 13]). While extensive form theory is based on graph theory and refined partitions, stochastic analysis describes exogenous information by means of filtrations on a given measurable space of scenarios, which are strictly less restrictive objects in that σ-algebras need not be generated by partitions of the sample space. This works as if a one-shot lottery draw selected a scenario ω in the beginning without being communicated and then, as time progresses, more and more properties of ω became known to the decision maker in question. This can be, for instance, the knowledge about the realised path of a Brownian motion. From a decision-theoretic point of view, there is no loss in working with this weaker structure as concerns exogenous information.

As to the actual decision makers, however, the stochastic analysis-inspired approach is not compatible with the extensive form paradigm in a strict sense, because it does not explain strategies, outcomes, (subgame-perfect) equilibria in terms of a decision tree-like object and choices locally available at moves. In this approach, typically, so-called "strategies" are defined as stochastic processes satisfying certain measurability property (e.g., progressive measurability) and such that a given stochastic differential equation depending on them in a non-anticipating way has a unique solution in some sense. This contradicts the problem's own extensive form characteristics in its strict sense because of the *ex post* restriction of the strategy space which would imply that the availability of a choice depends on future choices (see the introduction in [1] for a detailed discussion of this point). The "outcome" is taken to be the solution to the differential equation and "equilibria" (or "optimal controls") are defined with implicit reference to the dynamic programming principle, but without extensive form underpinning. Apart from fundamentally compromising the use of equilibrium concept as such, this lack can moreover cause some confusion, for instance, as to the decision-theoretic meaning of these "equilibria" (for example, concerning the notions of "closed" and "open loop" "equilibria" in stochastic differential games, see the discussion in [13, p. 72–76]) and to the precise definition of subgames (for example, in stochastic timing games, see [35] for a discussion).

Thus, there is scope for an extensive form theory that models exogenous information via filtration-like objects, but endogenous decision making via decision tree-like objects and local choices, therefore synthesising both approaches in a prolific way to what I call stochastic extensive forms. This is the aim of this project whose first part is this paper. This theory can be used to construct extensive form games with general stochastic processes as noise. What is more, it can be used to make precise in what sense stochastic control and differential games can be approximated by stochastic extensive form decision problems, for instance for the stochastic timing game, which is typically not formulated in extensive form (see in particular [35], where instead strategic form games are stacked according to a notion of "consistency" which is justified *a priori* by an analogy to discrete time and *a posteriori*, in [43], by a "discrete time with an infinitesimal grid" approximation argument reminiscent of [41, 42]). While a central motivation for this project stems from problems in continuous time, the theory is not restricted to this domain. For example, stochastic extensive form theory is independent of the representation of noise via a "nature" agent's virtual decision making, and in that sense truly stochastic.

The first step in this endeavour, which is taken in the present paper, is to build a theory of stochastic decision forests. These can be thought of as forests of rooted decision trees such that each tree corresponds to exactly one exogenous scenario, and equipped with a structure of similarity across trees identifying moves in a way that is consistent with the tree structure. Decision trees as graph-theoretical objects are the traditional base model for extensive form games and decision problems, but as pointed out in [3] their refined partitions-based representation in terms of the set of all maximal chains exhibits not only strong duality properties, but makes sequential and dynamic decision making amenable to the traditional decision-theoretic paradigm of choice under uncertainty. In this refined partitions approach, just as Savage's acts map states to outcomes ([38]), a strategy maps any move to a choice given by a set of outcomes still possible at this move from the agent's perspective. This is the way uncertainty about future choices (of oneself or opponents), that is, endogenous information is dealt with. The way uncertainty about exogenous noise alias exogenous information is modelled in this paper, however, is crucially different. It is the above-mentioned similarity structure, called random moves, that can be equipped with a filtration-like object dynamically revealing information about the realised exogenous scenario, allowing for general stochastic processes as drivers of noise without running into outcome generation problems along that dimension. The adaptedness of strategies can be based on a concept of adapted choices that is introduced in the final section.

The fundamental motivation for this theory arises in a large class of concrete applications. Thus, its presentation is tightly accompanied by examples. While some of them remain rather pedagogical, a general class of stochastic decision forests based on paths of action indexed by time is also constructed. A small set of readily verifiable axioms allows for many different time regimes, many different specifications of the outcomes (e.g. paths of timing games, also in the case of the scenario-dependent expiration of certain options), and general stochastic noise. Of course, it also includes the deterministic case. To the best of the author's knowledge, this theoretical unification of action path based decision problems with extensive form characteristics within one framework is a second new contribution of this paper in its own right.

Before ending the introduction, three remarks remain to be made. First, this is the first of three papers, hence, like in a piano sonata in three movements, the first part is self-contained and addresses an independent research problem in its own right, but it clearly prepares and is closely related to the other two papers and they must be thought together. Second, as pointed out by Aumann (see [6]), game theory arguably is "interactive decision theory", while decision theory is, in a trivial sense, single-player game theory. Still, game theory and decision theory are not the same, the first being more about the "interactive" and the latter more about the "decision". This paper is more concerned with the latter, hence it primarily employs the corresponding phraseology. But, if the context permits it, "game" will be used instead of "decision problem" or vice versa, and so on, and the reader is presumed to be able to deduce the meaning from the context. Finally, the relevant proofs of all new theorems, propositions, lemmata, and claims in examples, for all sections, can be found in the corresponding subsection of the appendix.

Notation

- $\mathbb{N} = \mathbb{Z}_+ =$ the natural numbers including zero, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, $\mathbb{R} =$ the real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$;

– a function $f\colon D \to V$ from a set D called *domain* to a set V is a subset of $D \times V$ such that for all $x \in D$ there is a unique $y \in V$ with $(x, y) \in f$, and this unique y is denoted by $f(x)$; in other words, f is described through its graph; as an abbreviation, the constant map on D with value $y \in V$ is denoted by $y_D = D \times \{y\}$;

– $\mathcal{P}(A) = \mathcal{P}A =$ the set of subsets of a given set A, $\mathcal{P}(f) = \mathcal{P}f =$ the function $\mathcal{P}A \to \mathcal{P}B$, $M \mapsto \{f(m) \mid m \in M\}$ for a given function $f\colon A \to B$ between two sets A and B;[1]

– $\operatorname{im} f = (\mathcal{P}f)(D) =$ the image of a set-theoretic function $f\colon D \to V$;

– $\bigcup M =$ the union of a set $M =$ the set of all x that are the element of some $S \in M$, also written $\bigcup_{i \in I} S_i$ in case M is the image of some function $I \ni i \mapsto S_i$, for some set I;

– $|M| =$ if M is not a number, then this is the cardinality of the set M;

– $> =$ the strict partial order associated to a weak partial order $\geq$;

– $\mathscr{A}|_D = \{A \cap D \mid A \in \mathscr{A}\}$, for any σ-algebra $\mathscr{A}$ and any $D \in \mathscr{A}$;

– $\mathscr{A}_1 \vee \mathscr{A}_2 =$ the smallest σ-algebra on Ω containing both $\mathscr{A}_1$ and $\mathscr{A}_2$, given a set Ω and sets $\mathscr{A}_1, \mathscr{A}_2 \subseteq \mathcal{P}(\Omega)$.

1. Decision forests

The basic object of classical extensive form decision and game theory is the decision tree. In the stochastic generalisation presented here, nature does not act as an agent taking decisions dynamically, but simply includes a device "randomly" selecting the decision tree the "personal" agents follow during their decision making process. This implies that we consider *decision forests*, rather than decision trees. Although this term is not new (as testified by [36]), we consider it in the context of abstract decision theory, and more precisely within the – decision-theoretically natural – refined partitions framework, and in the aim of making this framework amenable to exogenous noise in the sense of general probability theory. This framework has originally been developed for trees, rather than forests, pioneered in [45, section 8], and developed in much more generality in [3, 4] and subsequent papers, much of which is covered in the monograph [5]. In this section, we fix some order-theoretic language, present a definition and interpretation of decision forests within the refined partitions approach, and show that decision forests are nothing but forests of decision trees.

1.1. Order- and graph-theoretic conventions

We first recall some basic definitions from graph and order theory, thereby fixing conventions used in this text, which combine those from [12, 3, 16]. In a partially ordered set (in short, *poset*) $(N, \geq)$ a *chain* is a subset $c \subseteq N$ such that for all $x, y \in c$, $x \geq y$ or $y \geq x$ holds true. A *maximal chain* is a chain that is maximal as a chain with respect to set inclusion in $\mathcal{P}(N)$. $x \in N$ is called a *maximal element* iff there is no $y \in N$ other than x such that $y \geq x$. $x \in N$ is called *maximum* iff for all $y \in N$, $x \geq y$. For $x \in N$, the *principal up-set* $\uparrow x$ and *principal down-set* $\downarrow x$ are defined by

$$\uparrow x = \{y \in N \mid y \geq x\}, \qquad \downarrow x = \{y \in N \mid x \geq y\}.$$

[1] $\mathcal{P}$ defines a covariant endofunctor on the category of sets.

Moreover, in this text, a poset $(F, \geq)$ is called a *forest* iff for every $x \in F$, $\uparrow x$ is a chain. A forest $(F, \geq)$ is called *rooted* iff $F \neq \emptyset$ and for every $x \in F$, $\uparrow x$ contains a maximal element of $(F, \geq)$. A forest $(T, \geq)$ is called a *tree* iff for every $x, y \in T$, $(\uparrow x) \cap (\uparrow y) \neq \emptyset$. Given a forest $(F, \geq)$, the elements $x \in F$ are called *nodes*. Nodes $x \in F$ such that $\downarrow x = \{x\}$ are called *terminal*. We state the following lemma, fundamental for what follows. It can actually be seen as an explicitly order-theoretic reformulation of a basic result from graph theory (see the discussion in [12, section I.1]).[2]

Lemma 1.1. *For any forest $(F, \geq)$ there exists a unique partition $\mathcal{F}$ of F into trees such that for all $x, y \in F$ with $x \geq y$ there is $T \in \mathcal{F}$ with $x, y \in T$. If $(F, \geq)$ is rooted, then for any $T \in \mathcal{F}$, $(T, \geq)$ is a rooted tree and has a maximum.*

The elements of $\mathcal{F}$ are called *connected components* of $(F, \geq)$. The maximum of a rooted tree $(T, \geq)$ is called the *root*. The *roots* of a forest $(F, \geq)$ are the roots of its connected components. A *decision forest* (*decision tree*) is a rooted forest (tree, respectively) $(F, \geq)$ such that all $x, y \in F$ with $x \neq y$ can be separated by some maximal chain $c \subseteq F$, that is, $c \cap \{x, y\}$ is a singleton. A *move* in a decision forest $(F, \geq)$ is a non-terminal node $x \in F$.

If V is a set, a V*-poset* is a subset $N \subseteq \mathcal{P}(V)$. The name derives from the fact that $(N, \supseteq)$ defines a poset of subsets of V ordered by set inclusion.

Remark 1.2. Note that the definition of a tree used in this text is an order-theoretic transcription of a graph-theoretical concept. What is called forest here, corresponds indeed to a "forest", also named "acyclic graph" in graph theory (see [12, sections I.1, I.2]), but is called "tree" in [3, definition 1]. This latter terminology adapts the use of this term in order theory (see [16] and compare the discussion in [3, remark 2]). In the present text the use of forests with multiple connected components, describing exogenous scenarios, is central, and hence using the term "tree" for this may be misleading. Hence, the definitions of the present text insist stronger on the graph-theoretical aspect as presented in [12], and especially on the botanic metaphor of trees and forests, than do those in [3, 16], though without adding unnecessary discreteness assumptions.

Note, however, that a "rooted tree" in the sense of [3, definition 1] corresponds to a rooted tree in this text because the [3]-definition of "rooted" demands the existence of a maximum, not only of maximal elements for principal up-sets. But a rooted forest in the sense of this text need not be a "rooted tree" in the sense of [3].

It should also be noted that we use the partial order $\geq$ rather than $\leq$ because of the representation used in the subsequent subsection, following the decision- and game-theoretic texts [3, 5], though this may differ from the usual convention in other contexts.

1.2. Decision forests

The following definition is, formally, a transcription of the characterisation of (a subclass of) "game trees" in [3, theorem 3]. So, although, the object is formally not new, we look at it from a slightly different perspective which is why it is recalled here. First, it emphasises the fact that there can be multiple connected components, but restricts the attention to those [3]-"game trees" whose connected components are rooted. Second, the terminology in this text is different also in that it insists on the purely decision-theoretic aspect. What is called "decision forest (or tree) over a set" and "decision path" here, respectively, is called "game tree" and "play" in [3]. Third, while [3]-"game trees" are defined via certain set-theoretic properties which can be rather easily verified

[2] As the claim that it is a reformulation requires proof, and also for the reader's convenience, a proof can be found in the appendix.

in applications, and are then characterised via the so-called "representation by plays", the present text perceives the decision-theoretic essence of decision forests over sets rather as being the duality between outcomes and nodes expressed by that representation, and thus uses it as the definition.

Definition 1.3. Let V be a set. A *decision forest over* V is a V-poset F such that:

1. $(F, \supseteq)$ is a rooted forest;
2. F is *its own representation by decision paths*, that is, if W denotes the set of maximal chains in $(F, \supseteq)$, and for every $y \in F$, $W(y) = \{w \in W \mid y \in w\}$, then there is a bijection $f: V \to W$ such that for every $y \in F$, $(\mathcal{P}f)(y) = W(y)$.

F is called *decision tree over* V iff, in addition, for all $x, y \in F$ there is $z \in F$ with $z \supseteq x \cup y$.

The nodes, terminal nodes, and moves of $(F, \supseteq)$ are also called *nodes*, *terminal nodes*, and *moves* of F, respectively, and the elements of V are called *outcomes*. The set of moves of F is denoted by $X(F)$ or X in short.

Following [3], V can be seen as the set of possible *outcomes* of the game, and the forest $(F, \supseteq)$ specifies how, as the decision problem is tackled dynamically, the set of realisable outcomes becomes smaller and smaller. Thus, elements of F can be interpreted as nodes. See [3] for handy criteria on the set F characterising the situation of it being a decision forest over V: essentially, this corresponds to F being a "game tree" over V in the sense of [3, definition 4] such that all connected components of $(F, \supseteq)$ have maximal elements.

Condition (1) clarifies that the predecessors of any node $x \in F$ can be totally ordered and contain a root, and therefore constitute an account of the past with a beginning, a history. The set W of maximal chains can be interpreted as the set of *decision paths*. Condition (2) says that possible outcomes and decision paths are in one-to-one correspondence, in such a way that a node is contained in some decision path iff the corresponding outcome is contained in that node. Even more is true: according to [3, theorem 3, corollary 2]) and remark 1.2, the bijection f is uniquely determined. More precisely:

Proposition 1.4 ([3])**.** *Let F be a decision forest over some set V and $f: V \to W$ be a map as in definition 1.3, where W is the set of maximal chains in $(F, \supseteq)$. Then,*

$$\forall v \in V: \quad f(v) = \uparrow \{v\} = \{x \in F \mid v \in x\}.$$

Thus V, which formally is the set of possible outcomes, and W, which formally is the set of possible decision paths, can be identified in a uniquely determined way, namely via the f above. Under this identification, we have the following duality statement:

$$x \in f(v) \qquad \Longleftrightarrow \qquad v \in x,$$

for all nodes $x \in F$ and all outcomes $v \in V$. As a consequence, in the subsequent sections of the paper the set a decision forest is defined on is denoted by W rather than by V, consistent with other pieces of the literature, notably with [3, 5].

There is a second duality, namely between two approaches to dynamic decision theory: graphs, based on tree-like objects, and refined partitions, based on a set of outcomes. This is made precise in the following two propositions, which are contained in [3, lemma 14 and theorem 1], following remark 1.2:

Proposition 1.5 ([3])**.** *Let F be a decision forest over a set V. Then the poset $(F, \supseteq)$ is a decision forest.*

Proposition 1.6 ([3]). *Let $(F_0, \geq)$ be a decision forest. Let V be the set of its maximal chains and for any $x_0 \in F_0$ let $V(x_0)$ be the set of $v \in V$ with $x_0 \in v$. Let*

$$F = \{V(x_0) \mid x_0 \in F_0\}.$$

Then F defines a decision forest over V, and $(F, \supseteq)$ is order-isomorphic to $(F_0, \geq)$.

Rephrasing the "representation by plays" from [3], we call F the *representation by decision paths* of $(F_0, \geq)$. The definition of a decision forest over a set requires essentially that the operations from the two preceding proposition are, up to isomorphism, inverse to each other. See [3, section 4] for more details on this.

Thus, there is a rigorous sense in that both of the mentioned descriptions are equivalent. While the graph-theoretical approach is graphically more convenient, the refined partitions approach is in line with the classical decision theory of choice under uncertainty, in the spirit of Savage (see the introduction of [3] for a discussion). We note that, although phrased differently and with a different aim, all three preceding propositions and the duality concepts constitute one of the essential innovations of [3].

1.3. Forests of decision trees

In the present text, it is crucial to deal with forests and not only with trees, in the sense of subsection 1.1 and remark 1.2. Often, the study of forests can be reduced to the study of trees by considering the connected components separately. Is this true for decision forests as well, and in what sense? In other words, is the duality of nodes and outcomes compatible with the forest structure? This question gets fully answered in the following theorem. More precisely, we have:

Theorem 1.7. *Let V be a set and F be a V-poset defining a rooted forest with respect to $\supseteq$. Let $\mathcal{F}$ denote the set of connected components of the rooted forest $(F, \supseteq)$. For every $T \in \mathcal{F}$, let V_T denote the root of $(T, \supseteq)$.*

Then F is a decision forest over V iff

1. *$\{V_T \mid T \in \mathcal{F}\}$ is a partition of V, and*
2. *for every $T \in \mathcal{F}$, $(T, \supseteq)$ defines a decision tree over V_T.*

To verify the three propositions 1.4, 1.5, and 1.6, on can cite the results in [3] in combination with remark 1.2. In view of theorem 1.7, it suffices, however, to cite the results for the case of rooted trees only. This will be of interest at a later stage when analysing outcome existence and uniqueness for stochastic extensive forms.

Remark 1.8. Theorem 1.7 offers a way of constructing decision forests from collections of decision trees. Let $\mathcal{F}_0$ be a non-empty set of decision trees over sets. For any $T_0 \in \mathcal{F}_0$, let V_{T_0} be its root. Then, let

$$V = \bigcup_{T_0 \in \mathcal{F}_0} V_{T_0} \times \{T_0\} = \{(v, T_0) \mid T_0 \in \mathcal{F}_0, v \in T_0\},$$

the disjoint union of all roots, and let

$$F = \{x \times \{T_0\} \mid T_0 \in \mathcal{F}_0, x \in T_0\}.$$

Then, F is a decision forest over V. The set of connected components is given by

$$\mathcal{F} = \{\{x \times \{T_0\} \mid x \in T_0\} \mid T_0 \in \mathcal{F}_0\}.$$

Moreover, theorem 1.7 states that all decision forests over sets can be represented in this form.

For examples of decision forests over sets, the reader is referred to the following section, where decision forests whose connected components are indexed over the set of exogenous scenarios are considered.[3]

2. Stochastic decision forests

In this section, the central notion of this paper – stochastic decision forests – is introduced. The main idea behind this is to weaken the traditional assumptions on exogenous information, which is no more assumed to arise through the dynamic decision making of a nature agent. Rather, an exogenous scenario ω is "randomly" realised within a given measurable space which determines the decision tree underlying the actual decision makers' problem. Therefore, before the main definition, measurable spaces are discussed as a model for exogenous scenarios and sets of scenarios whose probability can be measured. Subsequently, the definition of stochastic decision forests is given and analysed, simple examples are presented, and the class of action path stochastic decision forests is constructed.

2.1. Exogenous scenario spaces

For the remainder of the paper, an *exogenous scenario space* is a measurable space $(\Omega, \mathscr{A})$ such that $\Omega \neq \emptyset$.[4] Ω describes the set of possible *exogenous scenarios* which replace the outcomes generated by the behaviour of a "nature" agent. $\mathscr{A}$ is the set of *events* alias collections of outcomes that can be measured by all relevant decision makers, in the sense, that their potential beliefs about events can be modelled via probability measures on $(\Omega, \mathscr{A})$. We formally follow the classical theory of subjective probability and its interpretation as a representation device of preferences in terms of expected utility, which goes back to Ramsey, de Finetti, Savage, among others. Accordingly, we think of an agent as an entity that has certain preferences over the result of the strategic interaction; according to the paradigm of the mentioned theory, this translates into a utility function over results and one or multiple probability measures alias beliefs on the exogenous scenario space.[5]

Why does one suppose $\mathscr{A}$ to be a σ-algebra? As an algebra of sets, $\mathscr{A}$ can be seen as a Boolean algebra and can therefore model basic logical operations on events. Moreover, in the probabilistic approach, we are only interested in those functions $\mathbb{P}\colon \mathscr{A} \to [0,1]$ compatible with the algebra structure[6] that are σ-additive, which can be thought of as a continuity property, and it is a fundamental measure-theoretic result, that, for such $\mathbb{P}$, without loss of generality, $\mathscr{A}$ can be supposed to be closed under countable unions, that is, to be a σ-algebra (see [11, section 1.5]).

If $\mathscr{A}$ can be generated by a countable partition $\mathscr{P}$ of Ω, any probability measure on it is entirely described through the probabilities on these partitioning events. So, though all elements of $\mathscr{A}$ including unions of partition members are events, one can formally describe the situation by the set $\mathscr{P}$ and a countable family of positive numbers $(p_A)_{A\in\mathscr{P}}$ adding up to 1 and indexed by that set. However, in general, $\mathscr{A}$ need not be generated by a partition of Ω. For instance, it may be that all singletons are events and have probability zero, as in the case of the Lebesgue measure on the unit interval $[0,1]$. Hence, from a general perspective, scenarios may be neither a relevant nor a sufficient description of exogenous data for decision makers, but it is the (σ-)algebra $\mathscr{A}$ of

[3] Also note the examples of decision trees over sets in [3] and the more extensive version in [5] in combination with the preceding remark 1.8.

[4] For more information on measure and probability theory, the reader is referred to the introductory texts [11, 28].

[5] See [22, chapter 2] for a rigorous mathematical textbook presentation of the theory with multiple priors, going back to [23], alias robust preferences, underpinning robust control and risk measures.

[6] The meaning of the phrase "compatible with the algebra structure" can be reduced to the statement: a) $\mathbb{P}(\Omega) = 1$, and b) for all disjoint $A_1, A_2 \in \mathscr{A}$ we have $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$.

events that these agents are concerned about. Moreover, as famously illustrated by the Lebesgue measure, in case Ω is uncountable, there is a non-trivial trade-off between the fineness[7] of $\mathscr{A}$ and the amount of probability measures on it (see [11, section 1.7]). Hence, to describe the agent's relation to exogenous data and admissible beliefs about their probability, it is crucial to specify $\mathscr{A}$. This is crucial, not the least because many relevant applications require an uncountable Ω (with a σ-algebra $\mathscr{A}$ that cannot be generated by a countable partition), be it the uniform distribution on $[0,1]$, be it continuous martingales omnipresent in finance (see, e.g., [17, 9]).

2.2. Stochastic decision forests

In light of the preceding argument, we model uncertainty about exogenous events by an exogenous scenario space $(\Omega, \mathscr{A})$ and suppose that some $\omega \in \Omega$ is realised, determining the relevant decision tree without being directly communicated to the decision makers. So, these agents do not necessarily know about the exogenous information telling in which tree they are while making choices based on the endogenous information given by moves, or "information sets" of moves – hence, there must be a structure of similarity among trees that can serve as a consistent basis for both endogenous and exogenous information revelation. This intuition is the basis of the following definition which is further interpreted and analysed in the sequel.

Definition 2.1. A *stochastic decision forest*, in short SDF, over an exogenous scenario space $(\Omega, \mathscr{A})$ is a triple $(F, \pi, \mathbf{X})$ consisting of:

1. a decision forest F over some set W;
2. a surjective map $\pi\colon F \to \Omega$ such that the set $\mathcal{F}$ of connected components of $(F, \supseteq)$ is given by the fibres of π, that is,
$$\mathcal{F} = \{\pi^{-1}(\{\omega\}) \mid \omega \in \Omega\};$$
3. a set $\mathbf{X}$ such that:
   (a) any element $\mathbf{x} \in \mathbf{X}$ is a section of moves defined on some event $D_{\mathbf{x}} \in \mathscr{A}$, that is, it is a map $\mathbf{x}\colon D_{\mathbf{x}} \to X$ satisfying $\pi \circ \mathbf{x} = \mathrm{id}_{D_{\mathbf{x}}}$;
   (b) $\mathbf{X}$ induces a covering of X, that is, $\{\mathbf{x}(\omega) \mid \mathbf{x} \in \mathbf{X},\ \omega \in D_{\mathbf{x}}\} = X$;
   (c) if we consider the partial order $\geq_{\mathbf{X}}$ on $\mathbf{X}$ given by
$$\mathbf{x}_1 \geq_{\mathbf{X}} \mathbf{x}_2 \quad \iff \quad \Big[D_{\mathbf{x}_1} \supseteq D_{\mathbf{x}_2} \ \text{ and } \ \forall \omega \in D_{\mathbf{x}_2} \colon \mathbf{x}_1(\omega) \supseteq \mathbf{x}_2(\omega)\Big],$$
   then for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}$:
$$\exists \omega \in D_{\mathbf{x}_1} \cap D_{\mathbf{x}_2} \colon \ \mathbf{x}_1(\omega) \supseteq \mathbf{x}_2(\omega) \qquad \Longrightarrow \qquad \mathbf{x}_1 \geq_{\mathbf{X}} \mathbf{x}_2;$$
   (d) for all $\mathbf{x} \in \mathbf{X}$, $\mathbf{x}(\omega)$ is a root for all or no $\omega \in D_{\mathbf{x}}$;
   (e) if $\bar{\mathbf{X}}$ is a set sharing the previous four properties of $\mathbf{X}$ that is *refined* by $\mathbf{X}$ in that for all $\bar{\mathbf{x}} \in \bar{\mathbf{X}}$ there is $P_{\bar{\mathbf{x}}} \subseteq \mathbf{X}$ with $\bar{\mathbf{x}} = \bigcup P_{\bar{\mathbf{x}}}$,[8] then $\bar{\mathbf{X}} = \mathbf{X}$;
   (f) for all $\mathbf{x} \in \mathbf{X}$, there is countable $\mathbf{X}_0 \subseteq \mathbf{X}$ such that for all $\mathbf{x}' \in \mathbf{X}$ with $\mathbf{x}' >_{\mathbf{X}} \mathbf{x}$ there is $\mathbf{x}_0 \in \mathbf{X}_0$ with $\mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}_0 \geq_{\mathbf{X}} \mathbf{x}$.

[7] Recall that a σ-algebra $\mathscr{A}'$ on Ω is *finer* than $\mathscr{A}$ iff $\mathscr{A} \subseteq \mathscr{A}'$.

[8] According to set-theoretic conventions, $\bar{\mathbf{x}} = \bigcup P_{\bar{\mathbf{x}}}$ means: $\bar{\mathbf{x}}$ is a map with domain $\bigcup_{\mathbf{x} \in P_{\bar{\mathbf{x}}}} D_{\mathbf{x}}$ and for all $\mathbf{x} \in P_{\bar{\mathbf{x}}}$ and $\omega \in D_{\mathbf{x}}$, $\bar{\mathbf{x}}(\omega) = \mathbf{x}(\omega)$.

The elements of $\mathbf{X}$ are called *random moves.* For $\omega \in \Omega$, let $T_\omega = \pi^{-1}(\{\omega\})$ and W_ω be the root for T_ω. For $A \subseteq \Omega$, let $W_A = \bigcup_{\omega\in A} W_\omega$ and $F_A = \bigcup_{\omega\in A} T_\omega$.

Let us discuss this definition in more detail. A stochastic decision forest is a decision forest over a set (1), whose connected components are indexed by Ω (2), admitting a maximal (3e) and separable (3f) set of sections of moves (3a) called random moves, that form a covering (3b) of X, and are compatible with the rooted forest $(F, \supseteq)$ by preserving the order (3c) and identifying roots (3d). As is going to be shown in the sequel, these properties are actually stronger than the definition formally indicates. If one visualises a stochastic decision forest in two dimensions, any tree growing along the vertical "decision path" axis, the forest's trees being placed along the horizontal "Ω" axis, then random moves are vertical sections partitioning the set of moves in a monotone, decision-theoretically interpretable way. Random moves thus define a structure of "similarity" across different connected components.

Moreover, they can be seen as moves of a rooted decision tree in their own right. Let us make this precise. The first thing to note is that as a consequence of the maximality assumption, random moves cannot be empty:

Lemma 2.2. *Let $(F, \pi, \mathbf{X})$ be a stochastic decision forest over an exogenous scenario space $(\Omega, \mathscr{A})$. Then, for all $\mathbf{x} \in \mathbf{X}$, $D_{\mathbf{x}} \neq \emptyset$.*

Yet, this is not sufficient to ensure that random moves can be distinguished as moves with respect to the order $\geq_{\mathbf{X}}$, since we might have to add terminal nodes, if they exist. For example, think of an "infinite centipede" which contains infinitely many maximal chains but only one maximal chain consisting only of moves. Hence, we extend the partial order $\geq_{\mathbf{X}}$ to a binary relation $\geq_{\mathbf{T}}$ on the disjoint union

$$\mathbf{T} = \mathbf{X} \cup \{(\omega, w) \in \Omega \times W \mid \{w\} \in F,\ \pi(\{w\}) = \omega\}$$

by letting, for $\mathbf{x} \in \mathbf{X}$ and $(\omega, w), (\omega', w') \in \mathbf{T} \setminus \mathbf{X}$:

- $\mathbf{x} \geq_{\mathbf{T}} (\omega, w)$ iff $\omega \in D_{\mathbf{x}}$ and $w \in \mathbf{x}(\omega)$;
- $(\omega, w) \geq_{\mathbf{T}} (\omega', w')$ iff $(\omega, w) = (\omega', w')$;
- $(\omega, w) \not\geq_{\mathbf{T}} \mathbf{x}$.

Note that a pair (ω, w) as above is nothing else than the set-theoretic function $\mathbf{y}\colon D_{\mathbf{y}} \to F, \omega \mapsto \{w\}$ where $D_{\mathbf{y}} = \{\omega\}$. In analogy with random moves and since its image is a terminal node, we call such $\mathbf{y}$ a *random terminal node*. The analogy is substantiated by the subsequent proposition which (seemingly) strengthens axiom 3c in the definition above.

Proposition 2.3. *Let $(F, \pi, \mathbf{X})$ be a stochastic decision forest over some exogenous scenario space $(\Omega, \mathscr{A})$ and $\mathbf{T}\bullet\Omega = \{(\mathbf{y}, \omega) \in \mathbf{T}\times\Omega \mid \omega \in D_{\mathbf{y}}\}$. Then the evaluation map $\mathrm{ev}\colon \mathbf{T}\bullet\Omega \to F, (\mathbf{y}, \omega) \mapsto \mathbf{y}(\omega)$ is a bijection such that for all $(\mathbf{y}_1, \omega_1), (\mathbf{y}_2, \omega_2) \in \mathbf{T} \bullet \Omega$:*

$$\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_2 \text{ and } \omega_1 = \omega_2 \quad \Longleftrightarrow \quad \mathbf{y}_1(\omega_1) \supseteq \mathbf{y}_2(\omega_2).$$

In order-theoretic terms, the evaluation map defines an order isomorphism between $T \bullet \Omega$, equipped with the order induced by the product of $\geq_{\mathbf{T}}$ on $\mathbf{T}$ and equality on Ω, and $(F, \supseteq)$. Now, as announced, we can make rigorous the sense in that random moves are the moves of a rooted decision tree.

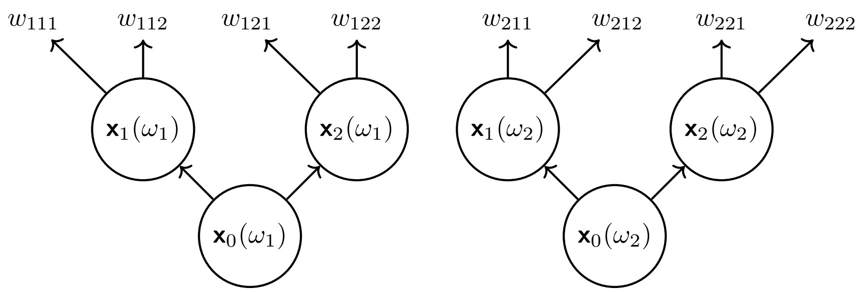

Figure 1: A simple stochastic decision forest represented as a directed graph, with $w_{\omega km} = (\omega, k, m)$, for $(\omega, k, m) \in W$; moves are indicated by circles

Theorem 2.4. *Let $(F, \pi, \mathbf{X})$ be a stochastic decision forest over some exogenous scenario space $(\Omega, \mathscr{A})$ such that all roots are moves.*[9] *Then, $(\mathbf{T}, \geq_{\mathbf{T}})$ defines a rooted decision tree and $\mathbf{X}$ is the set of its moves.*

How does an agent (alias decision maker) pass through a stochastic decision forest? There are actually two stories that can be told. On the one hand, as motivated in the outset of this subsection, first an exogenous scenario $\omega \in \Omega$ is determined, and then the agent passes through the decision tree given by the connected component $\pi^{-1}(\{\omega\})$. On the other hand, the agent passes through the rooted decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$. While the first perspective is the correct one in respect to outcome generation, the second one, though, is more relevant for the description of the exogenous information available to agents. The agents may not know ω or, even if they do, that knowledge may not be determinant.[10] The point of random moves lies in the fact that they serve as a flexible and general basis for "oracles", that is, a structure of exogenous information available to an agent. They will form the points that exogenous information is revealed at, and it is the tree property of $(\mathbf{T}, \geq_{\mathbf{T}})$ that will retain the extensive-form character of that process. In turn, the order-theoretic properties of a random move themselves do not reveal any exogenous information – this is a big difference to the classical approach to stochastic games and decision problems in extensive form, where exogenous information is the result of a "nature" player or agent making dynamic choices (going back to Shapley's work [39], see, e.g., [5, section 2.2.2.5]).

Note that stochastic decision forests are allowed to vary across scenarios: the trees in $\mathcal{F}$ need not be isomorphic. The crowns may become shallower in some scenarios (because some options are no more available there).

Remark 2.5. It can be checked in the proof in the appendix that proposition 2.3 holds true even if we do not assume properties 3d, 3e, and 3f.

2.3. Simple examples

In the following we present two very simple examples of stochastic decision forests. The first example is illustrated in figure 1. It indicates *pars pro toto* how finite stochastic extensive form decision problems can be formalised.

Let $\omega_1 = 1$, $\omega_2 = 2$ and consider $\Omega = \{\omega_1, \omega_2\}$, $\mathscr{A} = \mathcal{P}(\Omega)$, $W = \Omega \times \{1, 2\}^2$ and $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \colon \Omega \to \mathcal{P}(W)$ given by $\mathbf{x}_0(\omega) = \{\omega\} \times \{1, 2\}^2$ and $\mathbf{x}_k(\omega) = \{(\omega, k)\} \times \{1, 2\}$, $F = \{\mathbf{x}_k(\omega) \mid \omega \in \Omega,\ k =$

[9] Of course, any SDF can be reduced to this form by eliminating the connected components (alias exogenous scenarios) without moves.

[10] That point will be discussed in detail in subsection 3.1.

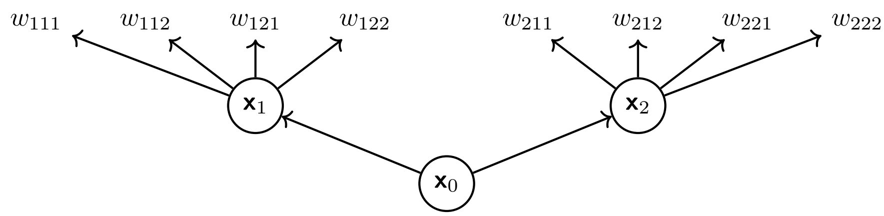

Figure 2: The decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$ for the simple stochastic decision forest, with $w_{\omega k m} = (\omega, k, m)$, for $(\omega, k, m) \in W$; (random) moves are indicated by circles

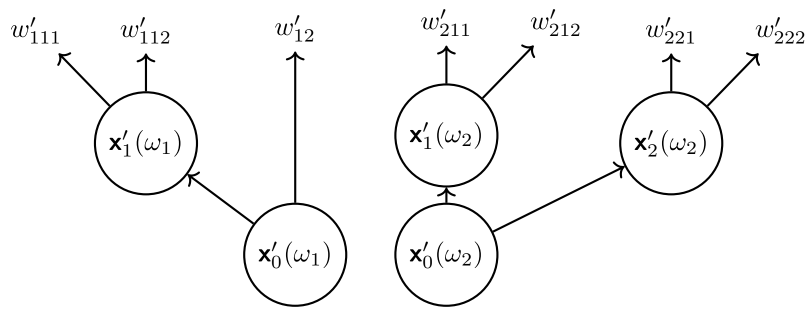

Figure 3: A variant of the simple stochastic decision forest in figure 1 represented as a directed graph, with $w'_{\omega k m} = (\omega, k, m)$, for all triples $(\omega, k, m) \in W'$, and $w'_{12} = (\omega_1, 2) = (1, 2)$; moves are indicated by circles

$0, 1, 2\} \cup \{\{w\} \mid w \in W\}$, $\pi\colon F \to \Omega$ be the map sending any node to the first entry of an arbitrary choice among its elements.

Lemma 2.6. *The tuple* $(F, \pi, \mathbf{X})$ *defines a stochastic decision forest.*

The corresponding decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$ is illustrated in figure 2.

As a variant, identifying the elements $(\omega_1, 2, 1)$ and $(\omega_1, 2, 2)$ in W will provide a stochastic decision forest with a random move that is not defined on all of Ω, as illustrated in figure 3. Put differently, let $W' = W \setminus \{(\omega_1, 2, 1), (\omega_1, 2, 2)\} \cup \{(\omega_1, 2)\}$ and let $\mathbf{x}'_0 = \mathbf{x}_0$, $\mathbf{x}'_1 = \mathbf{x}_1$, and $\mathbf{x}'_2\colon \{\omega_2\} \to \mathcal{P}(W')$ be given by $\mathbf{x}'_2(\omega_2) = \{(\omega_2, 2)\} \times \{1, 2\}$. Let $\mathbf{X}' = \{\mathbf{x}'_0, \mathbf{x}'_1, \mathbf{x}'_2\}$. Let $D_{\mathbf{x}'_0} = \Omega$, $D_{\mathbf{x}'_1} = \Omega$, $D_{\mathbf{x}'_2} = \{\omega_2\}$ and $F' = \{\mathbf{x}'(\omega) \mid \mathbf{x}' \in \mathbf{X}',\ \omega \in D_{\mathbf{x}'}\} \cup \{\{w'\} \mid w' \in W'\}$. Let $\pi\colon F' \to \Omega$ be the map sending any node to the first entry of an arbitrary choice among its elements.

Lemma 2.7. *The tuple* $(F', \pi', \mathbf{X}')$ *defines a stochastic decision forest.*

The corresponding decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$ is illustrated in figure 4.

2.4. Action path stochastic decision forests

In most pieces of the literature, dynamic games are defined by supposing a notion of time and specifying outcomes as certain paths of action at instants of time. [3, Subsection 2.2] provides a broad overview for this, including classical textbook definitions as in [21], infinite bilateral bargaining in discrete time as in [37], repeated games, the long cheap talk game in [7], and a decision-theoretic interpretation of differential games as in [18]. We desire to add, in a first approximate sense, stochastic control in both discrete and continuous time (see, e.g., [34, 10, 29]) and stochastic differential games (see, e.g., [13]) without restrictions on the noise in question.

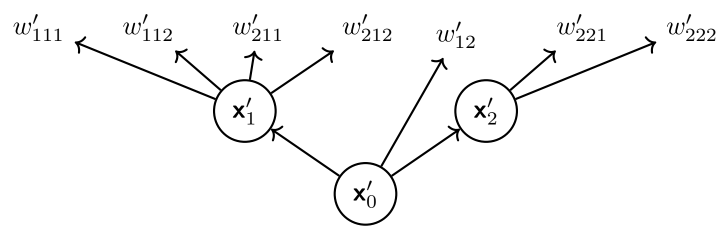

Figure 4: The decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$ for the variant of the simple stochastic decision forest, with $w'_{\omega km} = (\omega, k, m)$, for all triples $(\omega, k, m) \in W'$, and $w'_{12} = (\omega_1, 2) = (1, 2)$; (random) moves are indicated by circles

In this subsection, this approach is formulated in one abstract and general framework. This framework is based on a specific structure pertaining to all of these examples, namely *time*. Interestingly, time is not included in the abstract formulation of decision forests, and it serves as a particularly strong similarity structure for trees and even branches of one and the same tree. We do not address the question whether all (stochastic) decision forests can be represented in this framework. Of course, the second major point of the following framework is that it will allow for general exogenous stochastic noise, going strictly beyond the "nature" agent setting.

Let $\mathbb{T} \subseteq \mathbb{R}_+$ be such that $0 \in \mathbb{T}$ and $\mathbb{A}$ be a non-empty metric space. Let $(\Omega, \mathscr{A})$ be an exogenous scenario space. Let $W \subseteq \Omega \times \mathbb{A}^{\mathbb{T}}$ be such that for all $\omega \in \Omega$, there is $f \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, f) \in W$. An outcome will thus be a pair of an exogenous scenario and a path $f \colon \mathbb{T} \to \mathbb{A}$ of "action", and any scenario is required to admit at least one outcome. For any $t \in \mathbb{T}$ and $\tilde{w} = (\omega, f) \in \Omega \times \mathbb{A}^{\mathbb{T}}$, let

$$x_t(\tilde{w}) = x_t(\omega, f) = \{(\omega', f') \in W \mid \omega' = \omega,\ f'|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}}\}.$$

Let $F = \{x_t(w) \mid t \in \mathbb{T},\ w \in W\} \cup \{\{w\} \mid w \in W\}$. Further, let $\pi \colon F \to \Omega$ be the unique function mapping any $x \in F$ to the first item of one choice of its elements.

For $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$, let $D_{t,f} = \{\omega \in \Omega \mid |x_t(\omega, f)| \geq 2\}$. This will turn out as the event that $x_t(., f)$ is a move. We consider the following assumptions:

- **Assumption AP.W0.** For all $t \in \mathbb{T}$ and $f \in \mathbb{A}^{\mathbb{T}}$, $D_{t,f} \in \mathscr{A}$.
- **Assumption AP.W1.** For all $w \in W$ and all $t, u \in \mathbb{T}$ with $t \neq u$ and $x_t(w) = x_u(w)$, we have $x_t(w) = \{w\}$.
- **Assumption AP.W2.** For all $\omega \in \Omega$ and $\tilde{f} \in \mathbb{A}^{\mathbb{T}}$, and for all subsets $\mathbb{T}' \subseteq \mathbb{T}$ satisfying $x_t(\omega, \tilde{f}) \in F$ for all $t \in \mathbb{T}'$, there is $f \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, f) \in W$ and $f|_{[0,t)_{\mathbb{T}}} = \tilde{f}|_{[0,t)_{\mathbb{T}}}$ for all $t \in \mathbb{T}'$.
- **Assumption AP.W3.** For all $t \in \mathbb{T}$ and $f, g \in \mathbb{A}^{\mathbb{T}}$ such that $D_{t,f}, D_{t,g} \neq \emptyset$ and $D_{t,f} \cap D_{t,g} = \emptyset$, there is $u \in [0, t)_{\mathbb{T}}$ such that $D_{u,f} \cap D_{u,g} \neq \emptyset$ and $f|_{[0,u)_{\mathbb{T}}} \neq g|_{[0,u)_{\mathbb{T}}}$.
- **Assumption AP.W4.** There is a non-empty finite set I and a family of metric spaces $\mathbb{A}^i$, $i \in I$, such that $\mathbb{A} = \prod_{i \in I} \mathbb{A}^i$ in the category of topological spaces.

Assumption AP.W0 requires that $D_{t,f}$ is indeed an event in $(\Omega, \mathscr{A})$. AP.W1 stipulates that any (presumptive) move has a unique time associated to it. Indeed, we have the following result. For any $x \in F$, let

$$\mathbb{T}_x = \{t \in \mathbb{T} \mid \exists w \in x \colon x = x_t(w)\}.$$

It is evident that $\mathbb{T}_x$ is a non-empty convex set. We recall that convexity means that for all $t_0, t_1 \in \mathbb{T}_x$, we have $[t_0, t_1]_{\mathbb{T}} \subseteq \mathbb{T}_x$.

Lemma 2.8. *Assumption AP.W1 is satisfied iff for all $x \in F$ that are not singletons, $\mathbb{T}_x$ is a singleton.*

Note that the following weak converse statement holds true independently of whether AP.W1 is assumed or not:

Lemma 2.9. *All $x \in F$ such that $\mathbb{T}_x = \{t\}$ for some $t \in \mathbb{T}$ that is not maximal in $\mathbb{T}$, are no singletons.*

We immediately infer:

Corollary 2.10. *If assumption AP.W1 is satisfied and $\mathbb{T}$ has no maximum, then for all $x \in F$ the following statements are equivalent:*

– *x is a singleton.*

– *$\mathbb{T}_x$ is not a singleton.* □

Assumption AP.W2 corresponds to what is generally called "boundedness" in [3, subsection 3.4]. In the present context, this is the *conditio sine qua non* demanding that any (presumptive) decision path, alias maximal chain of nodes, corresponds to a path of action $\mathbb{T} \to \mathbb{A}$. Assumption AP.W3 refers to the maximality axiom 3e and ensures that different paths in different scenarios cannot be identified. If the assumption does not hold true, then it may be possible to identify the path f on $D_{t,f}$ with the path g on $D_{t,g}$, as for all $u < t$ until that (exclusively) f and g can be distinguished, $D_{u,f}$ and $D_{u,g}$ continue to be disjoint. Finally, assumption AP.W4 describes the situation where an action is the product of the individual actions of a finite set of I "agents". Of course, this assumption is trivially satisfied for singleton I and $\mathbb{A}^i = \mathbb{A}$ for the unique $i \in I$, but sometimes it is satisfied for larger I. Hence, the relevant point about AP.W4 is not whether it is satisfied, but with respect to what family $(\mathbb{A}^i)_{i \in I}$ it is.

Further, let $\mathbf{X}$ be the set of maps

$$\mathbf{x}_t(f) \colon D_{t,f} \to F, \omega \mapsto \mathbf{x}_t(f)(\omega) = x_t(\omega, f),$$

ranging over all $t \in \mathbb{T}$, $f \in \mathbb{A}^{\mathbb{T}}$ with $D_{t,f} \neq \emptyset$.

An important result of this paper, expressed in the next theorem, is that the preceding construction – encompassing a large part of decision problems and games, including very general stochastic versions of it – is well-defined and yields an SDF. This provides the basis for formulating a large class of stochastic decision problems in extensive form – already in this paper it is shown how exogenous information and choices adapted to it, and in the following papers how decision problems can be modelled on that basis.

Theorem 2.11. *Under assumptions AP.W0, AP.W1, AP.W2, and AP.W3, the triple $(F, \pi, \mathbf{X})$ from the current subsection defines a stochastic decision forest over $(\Omega, \mathscr{A})$.*

As we will discuss in the follow-up article, for well-ordered $\mathbb{T}$, well-posed decision problems can be defined on this action path SDF, while for the general case, this SDF can turn out to be ill-posed. For this latter case, the action path SDF represents an important first step though, in that a modification of it, the action-reaction SDF, a generalisation of a concept in [1], will restore well-posedness, as will be discussed in the third paper of the present series.

Definition 2.12. An *action path SDF* is a stochastic decision forest $(F, \pi, \mathbf{X})$ over an exogenous scenario space $(\Omega, \mathscr{A})$, that is constructed as above out of a set W satisfying assumptions AP.W0, AP.W1, AP.W2, and AP.W3. $(F, \pi, \mathbf{X})$ is said to be *induced* by W. $\mathbb{A}$ is called its *action space* and $\mathbb{T}$ its *time axis*. If W satisfies assumption AP.W4, then the elements of I are called *agents* and, for any agent $i \in I$, $\mathbb{A}^i$ is called the *action space for agent i*.

Given an action path SDF as above, let $\mathfrak{t}\colon X \to \mathbb{T}$ be the map assigning to any move $x \in X$ the unique element $\mathfrak{t}(x)$ of $\mathbb{T}_x$, see lemma 2.8.

Lemma 2.13. *Let $(F, \pi, \mathbf{X})$ be an action path SDF and $W = \bigcup F$. Then $\mathfrak{t}$ is strictly decreasing, that is, for all $x, y \in X$ with $x \supsetneq y$ we have $\mathfrak{t}(x) < \mathfrak{t}(y)$. Moreover, for all $\mathbf{x} \in \mathbf{X}$ and all $\omega, \omega' \in D_{\mathbf{x}}$, $\mathfrak{t}(\mathbf{x}(\omega)) = \mathfrak{t}(\mathbf{x}(\omega'))$.*

By this lemma, $\mathfrak{t}$ induces a map $\mathbf{X} \to \mathbb{T}$ which we denote also by $\mathfrak{t}$.

We conclude this section with examples of action path SDF. This includes the illustrative examples from subsection 2.3 and several typical classes of well-known decision problems. It remains an open question at this point whether any stochastic decision forest can be represented as an action path SDF, that is, so to speak, whether for any SDF all branches of all trees can be synchronised.

Lemma 2.14. *The two simple stochastic decision forests from subsection 2.3 can be represented as an action path forest with time $\mathbb{T} = \{0, 1\}$.*

The proof of this lemma in the appendix shows that the representation of the variant $(F', \pi', \mathbf{X}')$ as an action path SDF has to specify a "dummy" action at the terminal node $\{w'_{12}\}$, without letting an alternative, that is, an action meaning inaction. This is an artefact of the modelling decision to explicitly include a temporal dimension, as reflected in the action path formulation.[11] Stochastic decision forests are based on first principles and do no include such a dimension. Therefore, apart from being more general and flexible, their decision-theoretic interpretation is much clearer. At the same time, they include action path SDFs which model a structure present in many applications and which can therefore simplify the formal representation, at the cost of possibly introducing artificial phenomena like inactive activity. It may be noted that, in action path SDF, time plays a role complementary to that of random moves: it defines similarity of moves across branches, while random moves define similarity across trees. The random moves in action path SDFs are defined such as to be perfectly compatible with time.

Example 2.15.

- For $\mathbb{T} = \mathbb{R}_+$, singleton Ω, $W = \Omega \times \mathbb{A}^{\mathbb{T}}$, the action path stochastic decision forest yields the game tree of what is called "differential game" in [3, paragraph 2.2.5]. Hence, the action path SDF is a generalisation of this example in several directions: first and foremost, by adding an exogenous scenario space, second, by generalising the time axis, and third, by allowing for much higher flexibility regarding the set of outcomes.
- For $\mathbb{T} = \{0, \dots, N\}$, for $N \in \mathbb{N}^*$, or $\mathbb{T} = \mathbb{N}$ the corresponding action path SDF can serve as a basis describe the classical finite or infinite horizon discrete time, stochastic decision problem (see, e.g., [10]). The possible actions can be specified through the choice of W. The case $W = \Omega \times \mathbb{A}^{\mathbb{T}}$ corresponds to the case where at each move, the set of possible actions is $\mathbb{A}$. It is shown in the appendix that this W induces an action path SDF, even for general $\mathbb{T}$.

  Note that in contrast to the traditional "nature" model of stochastic games and decision problems in discrete time, exogenous information cannot be deduced from the order-theoretic properties of the (random) moves. The "nature" player or agent is replaced with a forest of decision trees and a structure of random moves to that we will later attach exogenous information in the form of σ-algebras.

[11] Assumption AP.W1 ensures that moves have a unique time associated to them, but terminal nodes need not. In many cases they do not, as made apparent by the lemmata 2.8 and 2.9 and, most strikingly, their corollary 2.10. But there may be an instant of time that certain branches still "move" at, and others do not and actually identify that instant with later points in time, which appears a bit artificial.

– $\mathbb{T}$ can also equal more general well-orders: For $\mathbb{T} = \{1 - 2^{-n} \mid n \in \mathbb{N}\} \cup \{1\}$, the long cheap talk game tree (see [7]) – and general stochastic variants thereof in the sense of SDFs – can be obtained. This example, in case of singleton Ω, is treated in [3].

– Suppose that assumption AP.W4 is satisfied with $\mathbb{A}^i = \{0, 1\}$ for all $i \in I$. Let W be the set of pairs (ω, f) where $f : \mathbb{T} \to \mathbb{A}$ is component-wise decreasing. It is shown in the appendix that W induces an action path SDF. The corresponding action path SDF is a natural (and at least approximate) candidate for describing stochastic decision problems of *timing* (alias *stopping*) (see, e.g., the monographs [40, 33, 19] for the mathematical theory, and [35, 29] for the link to applications in economics and finance).

– Let us consider an example of a forest that becomes shallower towards its crown in some areas of the exogenous scenario space. We use an example from finance, namely the exercise of an American up-and-out option ([27, chapter 26]). Let $P = (P_t)_{t \in \mathbb{R}_+}$ be a continuous stochastic process on $(\Omega, \mathscr{A})$ with strictly positive real values describing the price of a financial security. The initial price is $P_0 = 1$. As long as the price has not reached 2, the holder of the option can exercise it, but once the price reaches 2, the option expires irreversibly. This problem can be modelled using the action path stochastic decision forest associated to the following set of outcomes W.

  Take W to be the set of all (ω, f) where $\omega \in \Omega$ and $f : \mathbb{R}_+ \to \{0, 1\}$ is decreasing such that, if f takes the value 0, then $t^*_f = \inf\{t \in \mathbb{R}_+ \mid f(t) = 0\}$ satisfies $\max_{t \in [0, t^*_f]} P_t(\omega) < 2$. It is shown in the appendix that W induces an action path SDF, and that

$$D_{t,f} = \{\omega \in \Omega \mid \max_{u \in [0,t]} P_u(\omega) < 2\},$$

  for all $t \in \mathbb{R}_+$ and decreasing $f \in \{0, 1\}^{\mathbb{R}_+}$ with $f(t-) = 1$.

3. Exogenous information

As observed in the previous section, exogenous information is not contained in the order-theoretic structure of random moves. The approach put forward in this paper consists in attaching exogenous information to random moves in the form of σ-algebras, in a way analogous to (and in some sense, more general than) the use of filtrations in probability theory. Therefore, this latter approach is analysed and interpreted in the first subsection, and based on that, in the following subsections, a concept of exogenous information revelation on stochastic decision forests is introduced and explained, as well as illustrated by examples for the stochastic decision forests encountered in the previous section.

3.1. Filtrations in probability theory

Recall that for a measurable space $(\Omega, \mathscr{A})$, the usual interpretation of $\mathscr{A}$ sees its elements as being those subsets of Ω that are measurable for a given observer or agent. In the context of what we called exogenous scenario spaces, this had essentially the meaning that for all admissible beliefs (alias probability measures) on Ω, probabilities can be computed for these subsets (see subsection 2.1). However, there is a second meaning to "measurable", ubiquitous in probability theory and many of its applications.

In probability theory, a filtration is a monotone map $\mathscr{F}$ from a non-empty totally ordered set $\mathbb{T}$ modelling time, typically a suborder of the two-sided compactification of $\mathbb{R}$, to the set of sub-σ-algebras of $\mathscr{A}$, thus mapping any $t \in \mathbb{T}$ to a sub-σ-algebra $\mathscr{F}_t \subseteq \mathscr{A}$ such that for all $t, u \in \mathbb{T}$

with $t \leq u$, $\mathscr{F}_t \subseteq \mathscr{F}_u$ (see, e.g., [28, chapter 9] or [14, chapter 3]). The customary interpretation refers to an agent being equipped with that filtration $\mathscr{F}$ such that, for all times $t \in \mathbb{T}$, $\mathscr{F}_t$ describes the information the agent has at time t, or put differently, $\mathscr{F}_t$ is the set of events measurable for that agent at time t. Of course, this agent is able to compute probabilities for all $A \in \mathscr{A}$, but the elements of $\mathscr{F}_t$ are measurable in an even stronger sense.

Namely, one tacitly supposes that there is one realised scenario $\omega \in \Omega$, drawn "at random", so to speak, but information about it is only revealed to the agent progressively via $\mathscr{F}$, in the following way: For all instants of time $t \in \mathbb{T}$ and each $E \in \mathscr{F}_t$ the agent can discern at time t whether it contains the realised scenario or not. In particular, the agent when equipped with a belief in form of a probability measure on $\mathscr{A}$ can derive updated probabilities and expectations by evaluating conditional probabilities and expectations on E. Yet, this notion of measurement must be read with some caution. Actually, this measurement capacity of the agent includes logical operations given by the σ-algebra property, in particular countable logical disjunctions alias unions. Uncountable operations might be excluded however. Moreover, under the belief of the agent, E may have probability zero and hence the evaluation of conditional probabilities and expectations on E may be meaningless ([28, 14]). This can be illustrated by the following examples.

In the discrete setting, where $\mathscr{A}$ is generated by a countable partition of Ω, the interpretation is evident. Then all $\mathscr{F}_t$, $t \in \mathbb{T}$, are also generated by countable partitions, and the partition of later instants of time refine the preceding partitions. The members of the partition generating $\mathscr{F}_t$ can be thought of as the agent's information sets regarding the "nature" agent's past choices at time t.[12] In that discrete case, thus, filtrations are formally and conceptually a special case of order-theoretic rooted forests over the set Ω, as discussed in section 1.

But in general, $\mathscr{F}_t$, for some $t \in \mathbb{T}$, may, for instance, contain all the singletons subsets of Ω, without containing all elements of $\mathscr{A}$, let alone all subsets of Ω – think of Ω being the unit interval with Borel σ-algebra $\mathscr{A}$ and $\mathscr{F}_t$ being the set of subsets $A \subseteq [0, 1]$ that are countable or have countable complement. So the agent would actually be able to measure the realised scenario without discerning some $A \in \mathscr{A}$, say the interval $[0, \frac{1}{2}]$, although it is a union of measurable singletons. The point is that this union is uncountable, and we do not assume agents to be able to perform uncountable logical operations. Hence, the ability to discern the realised scenario does not necessarily imply the ability to discern all events, that is elements of $\mathscr{A}$. Moreover, all but countably many (and in the case of the Lebesgue measure on $[0, 1]$ even all) singletons must have probability zero, and so the ability to discern the realised scenario may not be relevant from a decision-theoretic perspective: If the conditional expectation of some interesting quantity, e.g., a payoff, given $\mathscr{F}_t$ is computed, it is completely inconclusive to evaluate it on some $E \in \mathscr{F}_t$ whose probability is zero, because the conditional expectation is defined almost surely.

Once again we note that the full descriptor of exogenous information at time t is the set of "measurable" events $\mathscr{F}_t$, where "measurable" is used in the second, stronger sense, not solely the set of scenarios Ω (compare the discussion in subsection 2.1). In particular, this discussion suggests that, in general, filtrations cannot be subsumed under the theory of decision forests and trees, which is "discrete" in that it can be represented via refined partitions, and thus filtrations provide a means to truly extend the scope of stochastic game and decision theory beyond those types of stochastic noise a dynamically choosing "nature" agent can simulate.

3.2. Exogenous information structures

Based on the preceding argument about exogenous scenario spaces and filtrations, and bearing in mind the way stochastic decision forests are built over exogenous scenario spaces, the following

[12] This will be discussed in detail in the follow-up paper.

definition is proposed, which is interpreted and analysed in the sequel.

Definition 3.1. Let $(F, \pi, \mathbf{X})$ be a stochastic decision forest over an exogenous scenario space $(\Omega, \mathscr{A})$. An *exogenous information structure* is a family $\mathscr{F} = (\mathscr{F}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$ such that for all $\mathbf{x} \in \mathbf{X}$:

1. $\mathscr{F}_{\mathbf{x}}$ is a sigma-algebra over $D_{\mathbf{x}}$ with $\mathscr{F}_{\mathbf{x}} \subseteq \mathscr{A}$;
2. for all $\mathbf{x}' \in \mathbf{X}$ with $\mathbf{x} \geq_{\mathbf{X}} \mathbf{x}'$ and every $E \in \mathscr{F}_{\mathbf{x}}$, we have $E \cap D_{\mathbf{x}'} \in \mathscr{F}_{\mathbf{x}'}$.

Concerning the sense of this definition, we think of any agent as being equipped with some exogenous information structure. For a random move $\mathbf{x} \in \mathbf{X}$, $\mathscr{F}_{\mathbf{x}}$ is interpreted as the set of events $E \in \mathscr{A}$ relevant at $\mathbf{x}$, meaning that $E \subseteq D_{\mathbf{x}}$, and that the agent can measure at $\mathbf{x}$, meaning that the agent can discern whether E contains the realised scenario or not. This second property is to be read in the sense discussed in the previous subsection 3.1. The more straightforward condition (2) ensures that the agent does not forget: if the agent is at $\mathbf{x}'$ and has been at $\mathbf{x}$ already, and was able to measure E at $\mathbf{x}$, then this agent can measure E given the domain of $\mathbf{x}'$ also at $\mathbf{x}'$. Note that, if for all $\mathbf{x} \in \mathbf{X}$, we have $D_{\mathbf{x}} = \Omega$, then an exogenous information structure is nothing but a monotone decreasing map $\mathscr{F} : \mathbf{x} \mapsto \mathscr{F}_{\mathbf{x}}$ from $\mathbf{X}$ into the set of sub-sigma-algebras of $\mathscr{A}$, "monotone decreasing" meaning that for all $\mathbf{x}, \mathbf{x}' \in \mathbf{X}$ with $\mathbf{x} \geq_{\mathbf{X}} \mathbf{x}'$, we have $\mathscr{F}_{\mathbf{x}} \subseteq \mathscr{F}_{\mathbf{x}'}$.

It should be noted that what we have at hand here is an adaptation of the notion of filtrations from probability theory to decision trees. In both cases, the index set is partially ordered. However, in the former case the order is total, whereas in the latter quite the opposite case is true, since it is a decision tree. Moreover, the underlying forest can "thin out" towards the crowns non-uniformly across trees alias exogenous scenarios, so that, in general, domains of random moves must be taken into account.

3.3. Simple examples

In the following we present examples of exogenous information structures for the simple examples from the previous section, subsection 2.3.

First, consider the simple stochastic decision forest $(F, \pi, \mathbf{X})$ from subsection 2.3, illustrated in figure 1. Consider the five following families $\mathscr{F} = (\mathscr{F}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$:

1. $\mathscr{F}_{\mathbf{x}} = \{\Omega, \emptyset\}$ for all $\mathbf{x} \in \mathbf{X}$: at all moves, it is unknown which scenario is realised;
2. $\mathscr{F}_{\mathbf{x}_0} = \{\Omega, \emptyset\}$ and one of the following three cases is true:
   (a) $\mathscr{F}_{\mathbf{x}_1} = \mathscr{F}_{\mathbf{x}_2} = \mathcal{P}(\Omega)$: only at the second move, it becomes known which scenario is realised, irrespective of which one is the second move;
   (b) $\mathscr{F}_{\mathbf{x}_1} = \mathcal{P}(\Omega)$, $\mathscr{F}_{\mathbf{x}_2} = \{\Omega, \emptyset\}$: $\mathbf{x}_1$ is the only move at that the realised scenario is revealed; an agent with this exogenous information may have interest in choosing (if possible) $\mathbf{x}_1$ rather than $\mathbf{x}_2$ in order to learn, modelling the trade-off *exploration vs. exploitation*; that way, problems with partial information and adaptive control can be modelled;
   (c) $\mathscr{F}_{\mathbf{x}_2} = \mathcal{P}(\Omega)$, $\mathscr{F}_{\mathbf{x}_1} = \{\Omega, \emptyset\}$: analogous to the preceding situation;
3. $\mathscr{F}_{\mathbf{x}} = \mathcal{P}(\Omega)$ for all $\mathbf{x} \in \mathbf{X}$: at all moves, the realised scenario is known.

Lemma 3.2. *For the simple stochastic decision forest $(F, \pi, \mathbf{X})$ from subsection 2.3 there are exactly five exogenous information structures, and these are given by the families $\mathscr{F}$ considered just above.*

Now, consider the variant $(F', \pi', \mathbf{X}')$ from subsection 2.3, as illustrated in figure 3. Consider the three following families $(\mathscr{F}'_{\mathbf{x}'})_{\mathbf{x}' \in \mathbf{X}'}$. In all three cases, let $\mathscr{F}'_{\mathbf{x}'_2} = \{\{\omega_2\}, \emptyset\}$. Moreover, separate the following three cases.

1. $\mathscr{F}'_{\mathbf{x}'_0} = \mathscr{F}'_{\mathbf{x}'_1} = \{\Omega, \emptyset\}$: again, there could be an exploitation vs. exploration trade-off (compare the cases 2(c) above);

2. $\mathscr{F}'_{\mathbf{x}'_0} = \{\Omega, \emptyset\}$, $\mathscr{F}'_{\mathbf{x}'_1} = \mathcal{P}(\Omega)$: this is similar to case 2(a) above;

3. $\mathscr{F}'_{\mathbf{x}'_0} = \mathscr{F}'_{\mathbf{x}'_1} = \mathcal{P}(\Omega)$: the realised scenario is known at any move (similar to case 3 above).

Lemma 3.3. *For the stochastic decision forest $(F', \pi', \mathbf{X}')$ from subsection 2.3 there are exactly three exogenous information structures, and these are given by the families $\mathscr{F}'$ considered just above.*

3.4. Action path stochastic decision forests

In this subsection, we discuss examples of exogenous information structures for the action path SDF $(F, \pi, \mathbf{X})$ induced by $W \subseteq \Omega \times \mathbb{A}^{\mathbb{T}}$ from subsection 2.4, defined over an exogenous scenario space $(\Omega, \mathscr{A})$. First, we observe a necessary condition on these: when evaluated along maximal chains of random moves in $(\mathbf{T}, \geq_{\mathbf{T}})$, we obtain a filtration in the sense of classical probability theory. More precisely:

Lemma 3.4. *Let $\mathscr{F}$ be an exogenous information structure for the action path SDF induced by a set W as in subsection 2.4, and let $f \in \mathbb{A}^{\mathbb{T}}$. Let $\mathbb{T}_f := \{t \in \mathbb{T} \mid D_{t,f} \neq \emptyset\}$.*

Then, for all $t, u \in \mathbb{T}_f$ with $t < u$ and $E \in \mathscr{F}_{\mathbf{x}_t(f)}$, $E \cap D_{u,f} \in \mathscr{F}_{\mathbf{x}_u(f)}$. In particular, if $D_{t,f} = \Omega$ for all $t \in \mathbb{T}_f$, then $(\mathscr{F}_{\mathbf{x}_t(f)})_{t \in \mathbb{T}_f}$ defines a filtration with time index set $\mathbb{T}_f$.

In the already discussed special case that $\mathbb{A}$ has at least two elements and $W = \Omega \times \mathbb{A}^{\mathbb{T}}$, $(\mathscr{F}_{\mathbf{x}_t(f)})_{t \in \mathbb{T}}$ is a filtration.

Now we turn to sufficient conditions, or more precisely, to the construction of exogenous information structures for the action path SDF. Let $\mathscr{G} = (\mathscr{G}_t)_{t \in \mathbb{T}}$ be a filtration on $(\Omega, \mathscr{A})$. For example, $\mathscr{G}$ could be generated by some stochastic process, possibly augmented with nullsets with respect to some set of probability measures on $(\Omega, \mathscr{A})$.

1. The basic setting in stochastic control of exogenous noise revealed over time independently of the agents' behaviour (see, e.g., [34] corresponds to letting, for all $\mathbf{x} \in \mathbf{X}$,

$$\mathscr{F}_{\mathbf{x}} = \mathscr{G}_{\mathsf{t}(\mathbf{x})} \mid_{D_{\mathbf{x}}},$$

and $\mathscr{F} = (\mathscr{F}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$.

2. We can also consider a more general case that allows for exogenous information depending on previous behaviour – which can serve as a basis for describing problems with partial information involving stochastic filtering, the trade-off exploration vs. exploitation, and adaptive control (see, e.g., [8, 14, 15]). Again, let $\mathscr{G} = (\mathscr{G}_t)_{t \in \mathbb{T}}$ be a filtration on $(\Omega, \mathscr{A})$. Let $(Y_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$ be a family of random variables with values in some metric space $\mathbb{B}$. Let, for any $\mathbf{x} \in \mathbf{X}$:

$$\mathscr{F}_{\mathbf{x}} = \Big(\sigma(Y_{\mathbf{x}'} \mid \mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}) \vee \mathscr{G}_{\mathsf{t}(\mathbf{x})}\Big) \Big|_{D_{\mathbf{x}}}.$$

Again, let $\mathscr{F} = (\mathscr{F}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$.

In particular, these two cases can be used to model "open loop" and "closed loop" controls, respectively, covering the definitions in [13, chapter 2], for example. However, as will be discussed in the section 4 about adapted choices, the crux lies in that local decisions not only depend on the filtration-like object representing exogenous information at the current random move, but also on the current random move itself. Hence, when comparing a counterfactual random move to a reference random move revealing the same exogenous information, the same strategy may demand a different choice, because the random move is different.

Example 3.5. We give a typical example of such a family $(Y_{\mathsf{x}})_{\mathsf{x}\in\mathbf{X}}$ in the case $\mathbb{B}=\mathbb{R}$, $\mathbb{A}=\mathbb{R}$, and $W=\Omega\times C(\mathbb{R})$. Consider an auxiliary family of real-valued random variables $(\tilde{Y}_{\mathsf{x}})_{\mathsf{x}\in\mathbf{X}}$ satisfying the stochastic differential equation

$$\tilde{Y}_{\mathsf{x}_t(f)} = \int_0^t \tilde{b}(f(u), \tilde{Y}_{\mathsf{x}_u(f)})\, \mathrm{d}\tilde{Z}_u,$$

for all $t\in\mathbb{T}$ and $f\in C(\mathbb{R})$, and bounded continuous $\tilde{b}\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ and a suitable $\mathscr{G}$-adapted stochastic integrator $\tilde{Z}$, say Brownian motion. Then let $(Y_{\mathsf{x}})_{\mathsf{x}\in\mathbf{X}}$ solve the stochastic differential equation

$$Y_{\mathsf{x}_t(f)} = \int_0^t b(\tilde{Y}_{\mathsf{x}_u(f)}, Y_{\mathsf{x}_u(f)})\, \mathrm{d}u + \int_0^t \sigma(Y_{\mathsf{x}_u(f)})\, \mathrm{d}Z_u$$

for all $t\in\mathbb{T}$ and $f\in C(\mathbb{R})$, and bounded continuous $b\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ and $\sigma\colon\mathbb{R}\to\mathbb{R}$, and a suitable $\mathscr{G}$-adapted stochastic integrator Z, say Brownian motion.

$\tilde{Y}$ can be interpreted as a noisy and not directly observable signal controlled by an agent through f, while Y can be seen as an observation depending on $\tilde{Y}$. This is a typical setting in control theory involving stochastic filtering (see, for instance, [8, 15], [14, part V]).

Theorem 3.6. *Consider the action path* SDF *induced by a set W, and the family $\mathscr{F}$ from above, in its general version as in point (2). Then, $\mathscr{F}$ defines an exogenous information structure on that* SDF*.*

4. Adapted choices

In this section, a concept of choices is introduced that aims at reconciling the classical decision-theoretic model of choice under uncertainty and the probabilistic setting of adapted processes. First, these two concepts, their differences and intersections are discussed. Then, the concept of adapted choices on stochastic decision forests is introduced and explained. The section will be completed with examples for the stochastic decision forests and exogenous information structures introduced in the two preceding sections.

4.1. Choice under uncertainty and adapted processes

In the refined partitions approach classical in sequential decision theory and developed in high generality in [3], a choice consists in selecting a member of a given partition of the set of outcomes that are still possible according to the information the given agent has at the current move. The partition members can be seen as “local outcomes”, pertaining to the current information. This is directly in line with the traditional approach of decision theory in the spirit of Savage ([38]): a strategy will then be nothing else than a Savage act selecting a local outcome at each move (alias “state” in Savage’s terms) in such a way that the set of possible outcomes is progressively reduced to a singleton.

In probabilistic models based on filtrations describing exogenous information on an abstract measurable space $(\Omega,\mathscr{A})$, local choices are modelled differently. If $\mathscr{F}=(\mathscr{F}_t)_{t\in\mathbb{T}}$ is a filtration on $(\Omega,\mathscr{A})$ with some totally ordered time index set $\mathbb{T}$, then the choice of an agent i equipped with that information and capable of action described by a metric space $\mathbb{A}^i$ at a time t is typically modelled by an $\mathscr{F}_t$-measurable function $g_t\colon \Omega \to \mathbb{A}^i$, modulo regularity conditions in t (leading to predictable processes, for instance), see, for instance, [34, 14]. $\mathscr{F}_t$-measurability of g_t means that for any measurable set of actions $A^i_t \subseteq \mathbb{A}^i$, the event that i chooses an action in A^i_t is measurable to agent i at time t, that is, $(g_t)^{-1}(A^i_t) \in \mathscr{F}_t$. Actually, A^i_t can be understood as a set-valued reference choice and $(g_t)^{-1}(A^i_t)$ is just the event that both choices, g_t and A^i_t, are compatible.

In the discrete case, this approach can be subsumed under the refined partitions framework going back to [45] and presented in high generality in [5]. Indeed, if $\mathscr{A}$ is generated by a countable partition, then so is $\mathscr{F}_t$ for each $t \in \mathbb{T}$, and $\mathscr{F}_t$-measurability is another way of expressing the requirement that g_t is constant on each member of the partition generating $\mathscr{F}_t$. Recall from subsection 3.1 that each realised partition member is thought of as the agent's information set regarding the "nature" agent's past choices. In that sense, g_t corresponds to a map from moves at time t to action such that at moves from the same information set the same action is selected. A family of such maps g_t, ranging over all times t, also called adapted process in the language of stochastic processes, then defines a complete contingent plan of action for this agent i provided i has only exogenous information.

On general measurable spaces, however, the $\mathscr{F}_t$-measurability cannot be rephrased like this in terms of partitions. Moreover, agents may have endogenous information about the own or other agents' past behaviour and condition their own behaviour on this information, that is, they may condition on what they know about their current random move. This suggests that one may adapt the more general measure-theoretic concept of measurable functions to stochastic decision forests as regards exogenous information about the horizontal Ω axis, while maintaining the refined partitions logic along the vertical tree axis in order to rigorously explain the interactive decision making in extensive form. The aim of this section is to introduce such a concept, bringing together the measure-theoretic probabilistic and the refined partitions based tree-like approaches to information.[13]

4.2. Choices in stochastic decision forests

A basic principle in extensive form theory is that at any "move" it is "known" to decision makers whether a given "choice" is available to them or not. One of the important facts the refined partitions approach formalised in [3] elucidates, is that the availability of a choice at a given move can be completely described in terms of the underlying set-theoretic structure: A choice is available at a move iff the latter is an immediate predecessor of the former. More precisely, if $(F, \pi, \mathbf{X})$ is a stochastic decision forest over an exogenous scenario space $(\Omega, \mathscr{A})$, $W = \bigcup F$ and $c \subseteq W$ is some subset (for instance, a union of nodes representing a choice), then, with

$$\downarrow c = \{x \in F \mid c \supseteq x\},$$

let, as in the classical setting of [3], the set of *immediate predecessors of* c be defined by:

$$P(c) = \{x \in F \mid \exists y \in \downarrow c \colon \uparrow x = \uparrow y \setminus \downarrow c\}.$$

Lemma 4.1. *For each $A \subset \Omega$, and each subset $c \subseteq W$,*

$$P(c \cap W_A) = P(c) \cap F_A.$$

[13] In this section we do however not make formally precise which axioms the set of choices an agent is equipped with should satisfy in order to consistently define a stochastic extensive form following the refined partitions paradigm and how exactly information sets can be modelled, nor do we make precise how and when this general stochastic approach can be represented using a "nature" agent: this is the aim of the follow-up paper.

As in [3], a *choice* is a non-empty union c of nodes. For $x \in X$, c is said to be *available at x* iff $x \in P(c)$, as in [3]. The sets $P(c)$ are a model for information sets in [3] and [4, 2] and will turn out to have, in a weaker sense, a similar role in the setting of stochastic decision forests, as discussed in the second paper.

While in a discrete setting the description of availability of choices in terms of predecessors makes it relatively easy to interpret and thus adhere to the above-mentioned basic principle, this becomes less evident in the present general measure-theoretic context, because of the tension between the discreteness of endogenous choices and the general measure-theoretic description of exogenous information in terms of systems of "measurable" sets rather than using partitions. We therefore explicitly assume a locally defined structure of reference choices like the A^i_t above, that can be measured by agents at this random move. Choices taken by agents must be adapted with respect to their individual exogenous information structure and that structure of reference choices. Hence, on the one hand, choices can be "discrete" in the sense of constituting partitions and thus eliminating a sufficient amount of alternatives so that progressive choosing results in a unique outcome. And on the other hand, choices can be adapted in the sense that the availability of its restriction to any reference choice is a measurable event with respect to the current exogenous information.

Definition 4.2. Let $(F, \pi, \mathbf{X})$ be a stochastic decision forest over an exogenous scenario space $(\Omega, \mathscr{A})$ and let $\mathscr{F}$ be an exogenous information structure.

1. A choice is said
   (a) *non-redundant* iff for any $\omega \in \Omega$ with $P(c) \cap T_\omega = \emptyset$, we have $c \cap W_\omega = \emptyset$;
   (b) *complete* iff for every random move $\mathbf{x} \in \mathbf{X}$, $\mathbf{x}^{-1}(P(c))$ is either empty or equal to $D_{\mathbf{x}}$.
2. For any random move $\mathbf{x} \in \mathbf{X}$, a non-redundant and complete choice c is said *available at* $\mathbf{x}$ iff $\mathbf{x}^{-1}(P(c)) = D_{\mathbf{x}}$.
3. A *reference choice structure* is a family $\mathscr{C} = (\mathscr{C}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$ of sets $\mathscr{C}_{\mathbf{x}}$ of non-redundant and complete choices available at $\mathbf{x}$.
4. Let $\mathscr{C}$ be a reference choice structure. An *$\mathscr{F}$-$\mathscr{C}$-adapted choice* is a non-redundant and complete choice c such that for all $\mathbf{x} \in \mathbf{X}$ that c is available at and all $c' \in \mathscr{C}_{\mathbf{x}}$:
$$\mathbf{x}^{-1}(P(c \cap c')) = \{\omega \in D_{\mathbf{x}} \mid \mathbf{x}(\omega) \in P(c \cap c')\} \in \mathscr{F}_{\mathbf{x}}.$$

From the agents' perspective, both for exogenous information revelation and adapted choices the relevant tree-like object is the decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$. In that respect, choices are made on that tree with respect to the exogenous information revealed along it. Regarding outcomes and outcome generation, however, the Ω dimension and thus the forest F are crucial. $\mathbf{X}$ builds the link between both, and exogenous information as well as adapted choices are defined with respect to $\mathbf{X}$ and such as to be compatible with each other. It is more general than usual continuous time stochastic control and differential games formulations (as in [34, 14, 13]) because of the dependence on $\mathbf{X}$, rather than on the more rigid notion of time. This point, among others, will be clarified in the following examples.

4.3. Simple examples

For the examples from subsection 2.3, both the basic version and its variation, we provide a reference choice structure, and a list of adapted choices, one for each exogenous information structure from example 3.3. Recall that $\Omega = \{1, 2\}$ and $\mathscr{A} = \mathcal{P}(\Omega)$.

First, consider the basic version $(F, \pi, \mathbf{X})$. Let M be the set of maps $\Omega \to \{1,2\}$. For $k \in \{1,2\}$ and $f, g \in M$, let

$$\begin{aligned} c_{f\bullet} &= \{(\omega, k', m') \in W \mid k' = f(\omega)\}, \\ c_{kg} &= \{(\omega, k', m') \in W \mid k' = k,\ m' = g(\omega)\}, \\ c_{\bullet g} &= \{(\omega, k', m') \in W \mid m' = g(\omega)\}. \end{aligned}$$

Define $c_{k\bullet}$, $c_{\bullet m}$, and c_{km} by identifying $k, m \in \{1,2\}$ with the constant maps on Ω with values k and m, respectively. Let $\mathscr{C}_{\mathbf{x}_0} = \{c_{1\bullet}, c_{2\bullet}\}$, $\mathscr{C}_{\mathbf{x}_1} = \mathscr{C}_{\mathbf{x}_2} = \{c_{\bullet 1}, c_{\bullet 2}\}$. Note the partitioned structure of these sets, reflecting the discreteness of the situation.

Lemma 4.3. *$\mathscr{C} = (\mathscr{C}_{\mathbf{x}})_{\mathbf{x} \in \mathbf{X}}$ defines a reference choice structure on $(F, \pi, \mathbf{X})$.*

Next, consider the following table. It reads as follows: Each line specifies a set of subsets of W for each of the five exogenous information structures (EIS) listed in subsection 3.3, first part; these subsets are classified according to whether they will correspond to choices at the beginning of the "first period" (at time 0) or of the "second period" (at time 1), if perceived as action path SDF according to lemma 2.14:

EIS	1st period	2nd period
1.	$c_{k\bullet} : k \in \{1,2\}$	$c_{km}, c_{\bullet m} : k, m \in \{1,2\}$
2.(a)	$c_{k\bullet} : k \in \{1,2\}$	$c_{kg}, c_{\bullet g} : k \in \{1,2\}, g \in M$
2.(b)	$c_{k\bullet} : k \in \{1,2\}$	$c_{1g}, c_{2m}, c_{\bullet m} : k, m \in \{1,2\}, g \in M$
2.(c)	$c_{k\bullet} : k \in \{1,2\}$	$c_{1m}, c_{2g}, c_{\bullet m} : k, m \in \{1,2\}, g \in M$
3.	$c_{f\bullet} : f \in M$	$c_{kg}, c_{\bullet g} : k \in \{1,2\}, g \in M$

Lemma 4.4. *For each exogenous information structure $\mathscr{F}$, the subsets of W given in the corresponding line of the preceding table are $\mathscr{F}$-$\mathscr{C}$-adapted choices on $(F, \pi, \mathbf{X})$.*

Second, consider the variant $(F', \pi', \mathbf{X}')$ of the simple stochastic decision forest. Let, again, M be the set of maps $\Omega \to \{1,2\}$. For $k \in \{1,2\}$ and $f, g \in M$, let

$$\begin{aligned} c'_{f\bullet} &= \{(\omega, k', m'), (\omega, k') \in W \mid k' = f(\omega)\}, \\ c'_{kg} &= \{(\omega, k', m') \in W \mid k' = k,\ m' = g(\omega)\}, \\ c'_{\bullet g} &= \{(\omega, k', m') \in W \mid m' = g(\omega)\}. \end{aligned}$$

Define $c'_{k\bullet}$, $c'_{\bullet m}$, and c'_{km} by identifying $k, m \in \{1,2\}$ with the constant maps on Ω with values k and m, respectively. Let $\mathscr{C}'_{\mathbf{x}'_0} = \{c'_{1\bullet}, c'_{2\bullet}\}$, $\mathscr{C}'_{\mathbf{x}'_1} = \mathscr{C}'_{\mathbf{x}'_2} = \{c'_{\bullet 1}, c'_{\bullet 2}\}$.

Lemma 4.5. *$\mathscr{C}' = (\mathscr{C}'_{\mathbf{x}'})_{\mathbf{x}' \in \mathbf{X}'}$ defines a reference choice structure on $(F', \pi', \mathbf{X}')$.*

Next, consider the following table. It reads as above.

EIS	1st period	2nd period
1.	$c'_{k\bullet} : k \in \{1,2\}$	$c'_{km}, c'_{\bullet m} : k, m \in \{1,2\}$
2.	$c'_{k\bullet} : k \in \{1,2\}$	$c'_{kg}, c'_{\bullet g} : k \in \{1,2\}, g \in M$
3.	$c'_{f\bullet} : f \in M$	$c'_{kg}, c'_{\bullet g} : k \in \{1,2\}, g \in M$

Lemma 4.6. *For each exogenous information structure $\mathscr{F}'$, the subsets of W' given in the corresponding line of the preceding table are $\mathscr{F}'$-$\mathscr{C}'$-adapted choices on $(F', \pi', \mathbf{X}')$.*

Note that the adaptedness of choices can be rephrased, namely by requiring that f and g be measurable with respect to the σ-algebra of exogenous information at the random move the choice is available at, respectively. This more convenient language is used in the context of action paths in the next subsection which contains the previous two examples following up on lemma 2.14.

Also note that a choice like c_{kg} is conditional on the knowledge that $\mathbf{x}_k$ has been chosen at the root of $\mathbf{X}$, while $c_{\bullet g}$ is independent of the initial choice. This is reflected by the respective sets of immediate predecessors, see lemma Appendix A.3. A similar remark is true for the variant, see lemma Appendix A.4. Hence, as in [3, chapter 5], choices can reflect the endogenous information, that is, the information about the position in the decision tree $(\mathbf{T}, \geq_{\mathbf{T}})$, agents have. While c_{kg} is a choice of perfect (endogenous) information, $c_{\bullet g}$ is not. The discussion on this theme will be continued in the second paper.

4.4. Action path stochastic decision forests

Finally, we consider the action-path SDF $(F, \pi, \mathbf{X})$ from subsection 2.4 over some exogenous scenario space $(\Omega, \mathscr{A})$. Let $W = \bigcup F$.

Let $t \in \mathbb{T}$. For any set $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ and any family $A_t = (A_{t,\omega})_{\omega\in\Omega}$ of Borel sets in $\mathbb{A}$, let

$$c(A_{<t}, A_t) = \{(\omega, f) \in W \mid f|_{[0,t)_{\mathbb{T}}} \in A_{<t},\ f(t) \in A_{t,\omega}\}.$$

For $t \in \mathbb{T}$, let $\mathscr{C}_t$ be the set of all $c(A_{<t}, A_t)$ ranging over all $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ and all families $A_t = (A_{t,\omega})_{\omega\in\Omega} \in \mathcal{P}(\mathbb{A})^{\Omega}$ of subsets of $\mathbb{A}$ satisfying the following assumptions:

– **Assumption AP.C0.** $c(A_{<t}, A_t) \neq \emptyset$.

– **Assumption AP.C1.** For all $w \in c(A_{<t}, A_t)$, there is $w' \in x_t(w) \setminus c(A_{<t}, A_t)$.

– **Assumption AP.C2.** For all $f \in \mathbb{A}^{\mathbb{T}}$ with $f|_{[0,t)_{\mathbb{T}}} \in A_{<t}$, we have

$$x_t(\omega, f) \cap c(A_{<t}, A_t) \neq \emptyset$$

for all or for no $\omega \in D_{t,f}$.

So we consider choices that correspond to actions at a predefined time t. Again, there is some sort of duality here: action paths are the result of progressive choosing; choices are collections of action paths, essentially. More precisely, $c(A_{<t}, A_t) \in \mathscr{C}_t$ describes the choice of an action in $A_{t,\omega}$ given scenario ω and at time t, given the history is contained in $A_{<t}$. We assume that such an action is really possible for at least some scenario (AP.C0), that it really constitutes a choices in that there is an alternative (AP.C1), and that it is complete in the sense that it only trivially intersects with random moves (AP.C2).

The principal down-sets and sets of predecessors of such choices take the expected form, as affirmed by the following lemmata. In particular, such a choice $c(A_{<t}, A_t)$ is available exactly at all those moves x, whose time $\mathfrak{t}(x)$ is t, and that contain an outcome compatible with the choice.

Lemma 4.7. *For all $t \in \mathbb{T}$ and $c \in \mathscr{C}_t$, we have:*

$$\downarrow c = \{x_u(w) \mid w \in c,\ u \in \mathbb{T} \colon t < u\} \cup \{\{w\} \mid w \in c\}.$$

Lemma 4.8. *For all $t \in \mathbb{T}$ and $c \in \mathscr{C}_t$, we have:*

$$P(c) = \{x_t(w) \mid w \in c\}.$$

As a result, we obtain a large class of non-redundant and complete choices for action path SDFs.

Lemma 4.9. *Let $t \in \mathbb{T}$ and $c \in \mathscr{C}_t$. Then, c defines a non-redundant and complete choice. Moreover, for all $\mathsf{x} \in \mathbf{X}$ that c is available at, there is $(\omega, f) \in c$ such that $\omega \in D_{t,f} = D_{\mathsf{x}}$ and $\mathsf{x} = \mathsf{x}_t(f)$.*

For the remainder of this subsection, we impose assumption AP.W4, see subsection 2.4. We fix both an agent $i \in I$ and an exogenous information structure $\mathscr{F}^i = (\mathscr{F}^i_{\mathsf{x}})_{\mathsf{x} \in \mathbf{X}}$ this agent is equipped with. We denote the canonical projection $\mathbb{A} \to \mathbb{A}^i$ by p^i. For all $t \in \mathbb{T}$ and $\mathsf{x} \in \mathbf{X}$ with $\mathsf{t}(\mathsf{x}) = t$, let $\mathscr{C}^i_{\mathsf{x}}$ the set of all sets $c(A_{<t}, A_t)$ as above such that

1. $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$;
2. $A_t = (A_{t,\omega})_{\omega \in \Omega}$ such that there is $A^i_t \in \mathscr{B}(\mathbb{A}^j)$ satisfying, for all $\omega \in \Omega$,
$$A_{t,\omega} = \begin{cases} (p^i)^{-1}(A^i_t), & \omega \in D_{\mathsf{x}}, \\ \emptyset, & \omega \notin D_{\mathsf{x}}; \end{cases}$$
3. $c(A_{<t}, A_t) \in \mathscr{C}_t$; and
4. for all $\omega \in D_{\mathsf{x}}$, $\mathsf{x}(\omega) \cap c(A_{<t}, A_t) \neq \emptyset$.

Hence, $c(A_{<t}, A_t)$ allows for choosing a measurable set of individual actions for agent i at the random move x.

Proposition 4.10. *$\mathscr{C}^i = (\mathscr{C}^i_{\mathsf{x}})_{\mathsf{x} \in \mathbf{X}}$ defines a reference choice structure.*

Let $t \in \mathbb{T}$, $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$, $D \in \mathscr{A}$, and $g \colon D \to \mathbb{A}^i$. Let $A^{i,g}_t = (A^{i,g}_{t,\omega})_{\omega \in \Omega}$ be given by
$$A^{i,g}_{t,\omega} = \begin{cases} \{a \in \mathbb{A} \mid p^i(a) = g(\omega)\}, & \omega \in D, \\ \emptyset, & \omega \notin D. \end{cases}$$

Let $c(A_{<t}, i, g) = c(A_{<t}, A^{i,g}_t)$. Provided it is an element of $\mathscr{C}_t$, this models the choice of agent i, given a history in $A_{<t}$, to take the action $g(\omega)$ in scenario $\omega \in D$ at time t. The perfect (endogenous) information case corresponds to $A_{<t} = \{f|_{[0,t)_{\mathbb{T}}}\}$ for some $f \in \mathbb{A}^{\mathbb{T}}$ such that $D_{t,f} \neq \emptyset$.

The next and final theorem of this paper is concerned with the following question: Provided $c = c(A_{<t}, i, g) \in \mathscr{C}_t$, what is the link between the adaptedness of c and the measurability of the function $g|_{D_{\mathsf{x}}}$ with respect to $\mathscr{F}^i_{\mathsf{x}}$, for all x that c is available at?[14] Indeed, in stochastic control and game theory, action path specifications are often implicitly used and choices are defined by assigning an action $g(\omega)$ to any scenario ω, for g measurable with respect to the value of the underlying filtration at the current time (see, e.g., [10, chapter 8], [34, 14, 13]). Hence, if we can answer the question above by providing a tight link, then this customary modelling paradigm in stochastic control and game theory can be explained as a derivative of adapted choices in stochastic decision forests and can therefore be interpreted in terms of traditional decision theory.

The theorem below affirms that the $\mathscr{F}^i_{\mathsf{x}}$-measurability of $g|_{D_{\mathsf{x}}}$ for relevant $\mathsf{x} \in \mathbf{X}$ is sufficient for the adaptedness of c. Moreover, it is also necessary, provided the reference choice structure sufficiently reflects the Borel σ-algebra on $\mathbb{A}^i$. Precisely, we consider the following assumption:

– **Assumption AP.C3.** For all $\mathsf{x} \in \mathbf{X}$ such that, with $t = \mathsf{t}(\mathsf{x})$, there are $D \in \mathscr{A}$, $g \colon D \to \mathbb{A}^i$ and $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ satisfying $\tilde{c} = c(A_{<t}, i, g) \in \mathscr{C}_t$ and making $\tilde{c}$ available at x, there are

[14] As stated in the theorem below, the availability of c at x implies $D_{\mathsf{x}} \subseteq D$.

1. $A'_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ such that for all $(\omega, f) \in \tilde{c}$, $f|_{[0,t)_\mathbb{T}} \in A'_{<t}$, and
2. an intersection-stable generator $\mathscr{G}(\mathbb{A}^i)$ of the Borel σ-algebra of $\mathbb{A}^i$ such that for all $G \in \mathscr{G}(\mathbb{A}^i)$, upon letting $A_t^{i,G} = (A_{t,\omega}^{i,G})_{\omega\in\Omega}$ be given by $A_{t,\omega}^{i,G} = (p^i)^{-1}(G)$ for $\omega \in D_\mathbf{x}$ and $A_{t,\omega}^{i,G} = \emptyset$ for $\omega \notin D_\mathbf{x}$, we have

$$c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}_\mathbf{x}^i \cup \{\emptyset\}.$$

Theorem 4.11. *Let $(F, \pi, \mathbf{X})$ be the action path* SDF *over $(\Omega, \mathscr{A})$ such that $W = \bigcup F$ satisfies assumption AP.W4. Further, let $i \in I$, $\mathscr{F}^i$ be an exogenous information structure and $\mathscr{C}^i$ be the reference choice structure defined above.*

Furthermore, let $t \in \mathbb{T}$, $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$, $D \in \mathscr{A}$, and $g\colon D \to \mathbb{A}^i$ such that $c(A_{<t}, i, g) \in \mathscr{C}_t$.

Then $c(A_{<t}, i, g)$ is a non-redundant and complete choice. Moreover, for all $\mathbf{x} \in \mathbf{X}$ that $c(A_{<t}, i, g)$ is available at, we have $D_\mathbf{x} \subseteq D$, and the following implications hold true:

1. *If $g|_{D_\mathbf{x}}$ is $\mathscr{F}_\mathbf{x}^i$-measurable, then $c(A_{<t}, i, g)$ is $\mathscr{F}^i$-$\mathscr{C}^i$-adapted.*
2. *If assumption AP.C3 is satisfied and $c(A_{<t}, i, g)$ is $\mathscr{F}^i$-$\mathscr{C}^i$-adapted, then $g|_{D_\mathbf{x}}$ is $\mathscr{F}_\mathbf{x}^i$-measurable.*

Example 4.12.

– For $\mathbb{T} = \mathbb{R}_+$, singleton Ω, singleton $A_{<t} = \{f|_{[0,t)_\mathbb{T}}\}$ for $t \in \mathbb{T}$, and $W = \Omega \times \mathbb{A}^\mathbb{T} \cong \mathbb{A}^\mathbb{T}$, $c(A_{<t}, i, g)$ corresponds to the so-called "differential game" choice in [2, section 2] (see also [5, example 4.14]), where agent i chooses the value of g given the history $f|_{[0,t)_\mathbb{T}}$. In that sense, the adapted choices constructed in theorem 4.11 generalise this example in several respects, most importantly along the stochastic dimension. If $\mathbb{T} = \{0, 1, \dots, N\}$ for some $N \in \mathbb{N}^*$ or $\mathbb{T} = \mathbb{N}$, then we get a model of the usual specification of choices from decision problems in discrete time (see, e.g., [10]).

– If we consider the simple SDFs from the preceding subsection in action path formulation, according to lemma 2.14, then the preceding construction yields exactly the adapted choices from the preceding subsection.

– If $W = \Omega \times \mathbb{A}^\mathbb{T}$, and $\mathbb{A}^i$ has at least two elements, and $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ is non-empty for some $t \in \mathbb{T}$, then for any map $g\colon \Omega \to \mathbb{A}^i$ we have $c(A_{<t}, i, g) \in \mathscr{C}_t$, that is, it satisfies assumptions AP.C0, AP.C1, and AP.C2. Moreover, assumption AP.C3 is satisfied.

– In case of the timing game (see example 2.15, and, e.g., [40, 29]), we have $\mathbb{A}^i = \{0, 1\}$. Let $t \in \mathbb{T}$. If $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ is a non-empty set of component-wise decreasing paths f_t such that $p^i \circ f_t = 1_{[0,t)_\mathbb{T}}$, then, for all $g\colon \Omega \to \{0, 1\}$, we have $c(A_{<t}, i, g) \in \mathscr{C}_t$, that is, it satisfies assumptions AP.C0, AP.C1, AP.C2. Moreover, assumption AP.C3 is satisfied.

– An analogous statement holds true in the case of the up-and-out option control problem (see example 2.15, and [27]). Here I is a singleton, $\mathbb{T} = \mathbb{R}_+$, and $\mathbb{A}^i = \mathbb{A} = \{0, 1\}$ for the unique $i \in I$. Let $t \in \mathbb{R}_+$, $A_{<t} = \{1\}^{[0,t)}$,

$$D = \{\omega \in \Omega \mid \max_{u\in[0,t]} P_u(\omega) < 2\},$$

and $g\colon D \to \{0, 1\}$ a map. If $D \neq \emptyset$, then we have $c(A_{<t}, i, g) \in \mathscr{C}_t$, that is, it satisfies assumptions AP.C0, AP.C1, AP.C2. Moreover, assumption AP.C3 is satisfied.

It is important to note that a choice $c = c(A_{<t}, i, g) \in \mathscr{C}_t$ as above, available at a random move $\mathbf{x}$, with $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ and $\mathscr{F}_\mathbf{x}^i$-measurable $g\colon D \to \mathbb{A}^i$, $D \in \mathscr{A}$, can condition on two distinct things: first, on the endogenous information $A_{<t}$, and second, on the exogenous information via

the $\mathscr{F}^i_{\mathbf{x}}$-measurability of g. For instance, one can the imagine the case in that there are $\mathbf{x}' \in \mathbf{X} \setminus \{\mathbf{x}\}$ with $\mathscr{F}^i_{\mathbf{x}} = \mathscr{F}^i_{\mathbf{x}'}$ and $\mathsf{t}(\mathbf{x}) = t = \mathsf{t}(\mathbf{x}')$, $\mathscr{F}^i_{\mathbf{x}}$-measurable $g' \colon D \to \mathbb{A}^i$ with $g \neq g'$ and $A'_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ such that $c(A'_{<t}, i, g')$ is available at $\mathbf{x}'$ and an agent's complete contingent plan of action specifies $c(A_{<t}, i, g)$ at $\mathbf{x}$, but $c(A'_{<t}, i, g')$ at $\mathbf{x}'$. In the context of example 3.5, with $\mathbf{x} = \mathbf{x}_t(f)$ and $\mathbf{x}' = \mathbf{x}_t(f')$ for suitable $f, f' \in \mathbb{A}^{\mathbb{T}}$, this would mean that, even if $Y_{\mathbf{x}_u(f)}$, $u \leq t$, and $Y_{\mathbf{x}_u(f')}$, $u \leq t$, generate the same σ-algebra, i.e., $\mathscr{F}^i_{\mathbf{x}_t(f)} = \mathscr{F}^i_{\mathbf{x}_t(f')}$, an agent need not make the very choice the agent makes at $\mathbf{x}_t(f)$ at the counterfactual random move $\mathbf{x}_t(f')$ as well.

Let us close on a note about this specific example for that we consider an agent only having choices of the form $c = c(A_{<t}, i, g)$ as in the preceding paragraph. In case $A_{<t} = \mathbb{A}^{[0,t)_{\mathbb{T}}}$ for all choices available to this agent, and moreover $\mathscr{F}^i_{\mathbf{x}_t(f)} = \mathscr{G}_t$ for all $f \in \mathbb{A}^{\mathbb{T}}$ and $t \in D_{t,f}$, and a fixed filtration $\mathscr{G} = (\mathscr{G}_t)_{t \in \mathbb{T}}$ (case 1 in subsection 3.4), then the complete contingent plans of action this agent can build, model what is described by the term "open loop" in the control- and game-theoretic literature. If, conversely, only choices with singleton $A_{<t}$ are available to the agent and $\mathscr{F}^i_{\mathbf{x}_t(f)}$ depends on a stochastic state controlled by f, for $f \in \mathbb{A}^{\mathbb{T}}$ and $t \in D_{t,f}$, in the sense of case 2 of subsection 3.4, possibly in combination with example 3.5, then we obtain what is subsumed under the term "closed loop" (see, e.g., [20, 21]).

Note that there are mixed regimes. "Large" alias uninformative $A_{<t}$ and $\mathscr{F}^i$ like in case 2 of subsection 3.4 may coexist as well, which is often also called "closed loop" in the control and differential games literature (see, e.g., the corresponding discussion in [13, p. 72–76]), but actually the complete contingent plans of action the agent can make only close the loop with respect to exogenous information: the agent can only react to the controlled noise, but not explicitly to the employed controls themselves which can make a crucial difference when evaluating counterfactuals, especially in case several agents interact. On the other hand, "small" alias informative $A_{<t}$ may coexist with case 1 of subsection 3.4, that is, "$\mathscr{F}^i_{\mathbf{x}_t(f)} = \mathscr{G}_t$", which is near, or in the extreme case of singleton $A_{<t}$ equal to closed loops with respect to endogenous information, but is equivalent, on the exogenous information side, to the ignorance of any controlled stochastic state. In any case, we conclude, that the terminology of "closed" and "open" loops must be used with respect to both endogenous and exogenous information. It requires some caution in the mixed regimes just discussed, especially if the underlying stochastic decision forest, exogenous information structures and adapted choices are not properly specified.

Conclusion

It is possible to implement general stochastic processes as background noise on refined partitions-based decision trees without running into outcome generation problems for a "nature" agent and allowing for a rigorous decision-theoretic interpretation of the relation of endogenous and exogenous information and choices. This is an improvement over the existing state of the art in both refined partitions-based decision and game theory in [5, 1], and stochastic control and differential games theory in [34, 29, 13, 14]. This has been achieved by abandoning the assumption of a "nature" agent, and instead constructing a theory of stochastic decision forests. These satisfy duality between outcomes and nodes and can be represented as forests of decision trees. They allow for a notion of similarity across moves on different trees, given by random moves, which necessarily form the moves of a decision tree in its own right. They serve as the basis for both a structure of exogenous information revelation similar to filtrations in probability theory, and a concept of adapted choices compatible with exogenous information structures which capture measurability assumptions on choices typically made in the literature and also shed light on the decision-theoretic meaning of open and closed loop controls in the stochastic setting.

In addition, a general model of action path stochastic decision forests has been constructed, given in terms of a small number of easily verifiable conditions. This model can serve as a unified decision-theoretic basis for a large class of stochastic decision problems, for instance in discrete time. Moreover, as will be shown in the third paper, it constitutes one step towards explaining stochastic decision problems in continuous time approximately.

In order to obtain a theory of stochastic extensive forms, encompassing the classical theory in [5], it still remains to formulate consistency criteria on stochastic decision forests equipped with a set of agents, each agent provided with an exogenous information structure, a reference choice structure and a set of adapted choices, to define strategies in the classical decision-theoretic sense going back to Savage, as prepared in this paper, and to ask the question whether and, if yes, how outcome existence and uniqueness (well-posedness) for strategy profiles can be characterised for stochastic extensive forms. Moreover, the question, implicit in the present paper, can then be formally addressed which stochastic extensive forms can be faithfully represented as classical extensive forms, that is, over a rooted decision tree, with a "nature" agent simulating the noise. An answer to this can formally substantiate the claim that stochastic extensive forms, and SDFs which they are based on, form a true and useful generalisation of the existing theory. Furthermore, we note in a spirit related to [35] that in order to allow for subgame-perfect equilibrium analysis, a decision-theoretically explainable and general model of subgames in stochastic extensive form decision problems needs to be established. All this constitutes the programme for the second paper of the present series.

Acknowledgements

First and foremost, I am particularly indebted to my doctoral supervisor Christoph Knochenhauer who read and commented on several earlier versions of this paper, encouraged me to pursue this project and actively supported me in presenting its contents at workshops. Special thanks are due to Frank Riedel for the helpful discussion we had in Berlin in 2022, especially for giving useful hints regarding the game-theoretic literature. I also want to thank Elizabeth Baldwin, Miguel Ballester, and Samuel N. Cohen for inspiring discussions on decision and control theory during my stay in Oxford in 2023. For their questions and comments, I thank the attendants of the workshops and seminars I had the opportunity to present and discuss earlier versions of this particular project at, in Oxford, Bielefeld, Kiel, Berlin, Palaiseau, including Daniel Andrei, Karolina Bassa, Peter E. Caines, Fanny Cartellier, Sebastian Ertel, Ali Lazrak, Kristoffer Lindensjö, Christopher Lorenz, Berenice Anne Neumann, Manos Perdikakis, Jan-Henrik Steg, Peter Tankov, and Jacco Thijssen. This project was partially funded by Deutsche Forschungsgemeinschaft (DFG) – Project ID 410208580 – IRTG2544 ("Stochastic Analysis in Interaction").

References

[1] Carlos Alós-Ferrer and Johannes Kern. Repeated games in continuous time as extensive form games. *Journal of Mathematical Economics*, 61:34–57, 2015.

[2] Carlos Alós-Ferrer, Johannes Kern, and Klaus Ritzberger. Comment on "Trees and extensive forms". *Journal of Economic Theory*, 146(5):2165—-2168, 2011.

[3] Carlos Alós-Ferrer and Klaus Ritzberger. Trees and decisions. *Economic Theory*, 25:764–798, 2005.

[4] Carlos Alós-Ferrer and Klaus Ritzberger. Trees and extensive form. *Journal of Economic Theory*, 143:216–250, 2008.

[5] Carlos Alós-Ferrer and Klaus Ritzberger. *The theory of extensive form games.* Springer, 2016.

[6] Robert J. Aumann. Game theory. In Matias Vernengo, Esteban Perez Caldentey, and Barkley J Rosser Jr, editors, *The New Palgrave dictionary of economics.* Palgrave Macmillan UK: Imprint: Palgrave Macmillan, 2020.

[7] Robert J. Aumann and Sergiu Hart. Long cheap talk. *Econometrica*, 71(6):1619–1660, 2003.

[8] Alan Bain and Dan Crisan. *Fundamentals of Stochastic Filtering.* Stochastic Modelling and Applied Probability. Springer, New York, 2009.

[9] Christian Bayer, Peter K. Friz, Masaaki Fukasawa, Jim Gatheral, Antoine Jacquier, and Mathieu Rosenbaum, editors. *Rough Volatility.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2023.

[10] Dimitri Bertsekas and Steven E. Shreve. *Stochastic optimal control: the discrete-time case*, volume 5. Athena Scientific, 1996.

[11] Vladimir I. Bogachev. *Measure theory*, volume 1. Springer, Berlin, Heidelberg, 2007.

[12] Béla Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics.* Springer Science & Business Media, New York, 2013.

[13] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications I. Mean field FBSDEs, control, and games*, volume 83 of *Probability theory and stochastic modelling.* Springer, Cham, 2018.

[14] Samuel N. Cohen and Robert J. Elliott. *Stochastic calculus and applications*, volume 2. Springer, 2015.

[15] Samuel N. Cohen, Christoph Knochenhauer, and Alexander Merkel. Optimal adaptive control with separable drift uncertainty. arXiv2309.07091, 2023. Preprint.

[16] Brian A. Davey and Hilary A. Priestley. *Introduction to lattices and order.* Cambridge University Press, 2nd edition, 2002.

[17] Freddy Delbaen and Walter Schachermayer. *The mathematics of arbitrage.* Springer, 2006.

[18] Engelbert J. Dockner, Steffen Jorgensen, Ngo Van Long, and Gerhard Sorger. *Differential Games in Economics and Management Science.* Cambridge University Press, 2000.

[19] Nicole El Karoui. Les aspects probabilistes du controle stochastique. In P. L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour IX-1979*, pages 73–238, Berlin, Heidelberg, 1981. Springer.

[20] Drew Fudenberg and David K. Levine. Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory*, 44(1):1–18, 1988.

[21] Drew Fudenberg and Jean Tirole. *Game Theory.* MIT Press, Cambridge (Massachusetts), London, 1991.

[22] Hans Föllmer and Alexander Schied. *Stochastic Finance.* De Gruyter, Berlin, Boston, 2016.

[23] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[24] Steven R. Grenadier. The strategic exercise of options: development cascades and overbuilding in real estate markets. *J. Finance*, 51(5):1653–1679, 1996.

[25] Xin Guo, Anran Hu, and Yufei Zhang. Reinforcement learning for linear-convex models with jumps via stability analysis of feedback controls. *SIAM Journal on Control and Optimization*, 61(2):755–787, 2023.

[26] Minyi Huang, Rolang P. Malhamé, and Peter E. Caines. Large population stochastic dynamic games. Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communication in information and systems*, 6(3):221–252, 2006.

[27] John C. Hull and Sankarshan Basu. *Options, futures, and other derivatives.* Pearson Education India, 10th edition, 2018.

[28] Olav Kallenberg. *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling.* Springer Cham, 2021. 3rd edition.

[29] Ioannis Karatzas and Steven E. Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.

[30] Harold W. Kuhn. Extensive games. *Proceedings of the National Academy of Sciences*, 36(10):570–576, 1950.

[31] Harold W. Kuhn. Extensive games and the problem of information. *Contributions to the Theory of Games*, 2(28):193–216, 1953.

[32] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2:227–260, 2007.

[33] Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems.* Birkhäuser, Basel, 2006.

[34] Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[35] Frank Riedel and Jan-Henrik Steg. Subgame-perfect equilibria in stochastic timing games. *Journal of Mathematical Economics*, 72:36–50, 2017.

[36] Lior Rokach. Decision forest: Twenty years of research. *Information Fusion*, 27:111–125, 2016.

[37] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109, 1982.

[38] Leonard J. Savage. *The foundations of statistics.* Dover Publications, New York, 1972.

[39] Lloyd S. Shapley. Stochastic games. *Proceedings of the national academy of sciences*, 39(10):1095–1100, 1953.

[40] Albert N. Shiryaev. *Optimal stopping rules*, volume 8. Springer Science & Business Media, 2007.

[41] Leo K. Simon. Basic timing games. Working paper 8745, University of California, Berkeley, 1987.

[42] Leo K. Simon and Maxwell B. Stinchcombe. Extensive form games in continuous time: Pure strategies. *Econometrica: Journal of the Econometric Society*, pages 1171–1214, 1989.

[43] Jan-Henrik Steg. On preemption in discrete and continuous time. *Dynamic games and applications*, 8(4):918–938, 2018.

[44] Maxwell B. Stinchcombe. Maximal strategy sets for continuous-time game theory. *Journal of Economic Theory*, 56(2):235–265, 1992.

[45] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

Appendix A. Proofs

Appendix A.1. Section 1

We first prove lemma 1.1 and theorem 1.7, without using the propositions 1.4, 1.5, and 1.6.

Proof of lemma 1.1. [15] For $x, y \in F$, let $x \sim y$ iff there is $z \in F$ such that $z \geq x$ and $z \geq y$. $\sim$ is clearly reflexive and symmetric. It is also transitive, because any principal up-set in $(F, \geq)$ is a chain and $\geq$ is transitive. The equivalence classes of $\sim$ define trees by construction of $\sim$. If there are $x, y \in F$ such that $x \geq y$, then (by taking $z = x$) we obtain that $x \sim y$. Hence, the equivalence classes of $\sim$ define a partition with the claimed property.

Let $\mathcal{F}$ be an arbitrary partition with the claimed property. Then, for $x, y \in F$, x and y belong to the same partition member iff $x \sim y$. Indeed, if they belong to the same partition member T, which by assumption is a tree, then $x \sim y$ by definition of a tree. If conversely $x \sim y$, then there is $z \in F$ with $z \geq x$ and $z \geq y$. Hence, there are $T, T' \in \mathcal{F}$ such that $x, z \in T$ and $y, z \in T'$. Thus $T \cap T' \neq \emptyset$, whence $T = T'$. Hence, $\mathcal{F}$ is uniquely determined.

We now prove the statement of the second sentence. Suppose the forest $(F, \geq)$ to be rooted, and let $T \in \mathcal{F}$. By definition of $\mathcal{F}$, T is non-empty. Let $x \in T$. Then the principal up-set $\uparrow x = \{y \in F \mid y \geq x\}$ in F contains a maximal element y. By definition of T, $y \in \uparrow x$ implies $y \in T$. Futhermore, let $z \in T$ with $z \geq y$. As $z \in F$, the maximality of y implies $z = y$. Hence, y is maximal in $(T, \geq)$ as well. Therefore, $(T, \geq)$ is rooted according to our definition.

It remains to prove the existence of a maximum of $(T, \geq)$. T is non-empty, hence there is $x \in T$. By assumption, there is a maximal element $y \in \uparrow x$. $\uparrow x$ is a chain, because $(F, \geq)$ is a forest. Hence, y is a maximum of $\uparrow x$ with respect to the induced order. Let $z \in T$ be an arbitrary element of T. Then $\uparrow x \cap \uparrow z \neq \emptyset$. Take an element u in this intersection, which is also an element of T, because $u \sim x$. Then $y \geq u \geq z$, whence $y \geq z$. Thus y is a maximum of $(T, \geq)$. □

We prepare the proof of theorem 1.7 with two lemmata.

Lemma Appendix A.1. *Let V be a set and F be a V-poset such that $(F, \supseteq)$ is a rooted forest. Let $\mathcal{F}$ be the set of its connected components and for any $T \in \mathcal{F}$, let V_T be the root of $(T, \supseteq)$. Then, for any $T \in \mathcal{F}$, we have:*

1. $T = \{x \in F \mid V_T \supseteq x\}$;
2. $W(V_T) = \{w \in W \mid w \subseteq T\}$;

[15] We recall that this lemma can actually be seen as an explicitly order-theoretic reformulation of a basic result from graph theory (see the discussion in [12, section I.1]). As the claim that it is a reformulation requires proof, and also for the reader's convenience, a proof is given nonetheless.

3. *for any $x \in T$, $W(x)$ is equal to the set of maximal chains w in $(T, \supseteq)$ with $x \in w$; in particular, $W(V_T)$ is equal to the set of maximal chains in $(T, \supseteq)$.*

Proof. (Ad 1): If $x \in T$, then $V_T \supseteq x$, because V_T is the root of T. Conversely, let $x \in F$ such that $V_T \supseteq x$. In combination with $V_T \in T$, we obtain $x \in T$, because $T \in \mathcal{F}$.

(Ad 2): Let $w \in W(V_T)$ and $x \in w$, then $V_T \supseteq x$ or $x \supseteq V_T$, whence $x \in T$, by definition of V_T and $\mathcal{F}$. Thus $w \subseteq T$. Conversely, let $w \in W$ such that $w \subseteq T$. By definition of V_T, any $x \in w$ satisfies $V_T \supseteq x$. Because of maximality of the chain w, we must have $V_T \in w$.

(Ad 3): Let $x \in T$.

Let $w \in W(x)$. Then $w \subseteq T$ since, for any $y \in w$ we have $x \supseteq y$ or $y \supseteq x$, whence $y \in T$. The chain w is also maximal in $(T, \supseteq)$ because T is a subset of F and w is maximal in $(F, \supseteq)$ by assumption.

Now let w be a maximal chain in $(T, \supseteq)$ with $x \in w$. By definition of V_T, we have $V_T \supseteq y$ for all $y \in w$, and maximality of the chain w implies that $V_T \in w$. Let $y \in F$ be such that $w \cup \{y\}$ is a chain in $(F, \supseteq)$. Hence, $V_T \supseteq y$ or $y \supseteq V_T$, in particular we get $y \in T$. As w is maximal in $(T, \supseteq)$, $y \in w$. Hence w is already maximal in $(F, \supseteq)$.

The second part follows from the first, using that all maximal chains w in $(T, \supseteq)$ satisfy $V_T \in w$. Indeed, as a maximal chain w is non-empty, and any element $x \in w \subseteq T$ satisfies $V_T \supseteq x$ by definition of V_T. By maximality of the chain w, we infer $V_T \in w$. □

Lemma Appendix A.2. *Let F be a decision forest over a set V with f as in definition 1.3. Let $\mathcal{F}$ be the set of $(F, \supseteq)$'s connected components. Then, for any $v \in V$ and any $T \in \mathcal{F}$, we have $f(v) \subseteq T$ iff $v \in V_T$.*

Proof. Let $v \in V$ and $T \in \mathcal{F}$. By proposition 1.4, $v \in V_T$ iff $V_T \in f(v)$. By lemma Appendix A.1, part (2), this is equivalent to $f(v) \subseteq T$. □

Proof of theorem 1.7. (Ad necessity): Suppose F to be a decision forest over V. Let W be the set of maximal chains in $(F, \supseteq)$, and $f\colon V \to W$ a bijection such that, for every $x \in F$, $(\mathcal{P}f)(x) = W(x)$, just as in definition 1.3.

We claim that for all $v \in V$ there is a unique $T \in \mathcal{F}$ such that $f(v) \subseteq T$. The claim (1) then follows with lemma Appendix A.2.

For the proof, let $v \in V$. As the maximal chain $f(v)$ is non-empty, there is $x \in f(v)$. As $\mathcal{F}$ is a partition of F, there is a unique $T \in \mathcal{F}$ with $x \in T$. For arbitrary $y \in f(v)$, we have $x \subseteq y$ or $y \subseteq x$, since $f(v)$ is a chain. Hence, $x \sim y$, or in other words $y \in T$. Thus T is the unique element of $\mathcal{F}$ with $f(v) \subseteq T$.

For claim (2), let $T \in \mathcal{F}$. Then $(T, \supseteq)$ defines a rooted tree with root V_T, by definition of $\mathcal{F}$. Moreover, for all $v \in V$, we have $f(v) \in W(V_T)$ iff $v \in V_T$, by proposition 1.4. Hence, $f_T := f|_{V_T}$ yields a bijection $V_T \to W(V_T)$.

By lemma Appendix A.1, part (3), second sentence, $W(V_T)$ equals the set of maximal chains in $(T, \supseteq)$. By definition of V_T, all $x \in T$ satisfy $V_T \supseteq x$ so that the of definition of f yields:

$$(\mathcal{P}f|_T)(x) = (\mathcal{P}f)(x) = W(x).$$

By lemma Appendix A.1, part (3), first sentence, $W(x)$ consists exactly of all maximal chains in $(T, \supseteq)$ containing x. Hence, T is its own representation by decision paths.

(Ad sufficiency): Let F satisfy properties (1) and (2) from the theorem. It remains to show that F is its own representation by decision paths, i.e., part (2) in definition 1.3.

By assumption (2) and part (3) of lemma Appendix A.1, for each $T \in \mathcal{F}$, there is a bijection $f_T\colon V_T \to W(V_T)$ such that for all $x \in T$, $(\mathcal{P}f_T)(x)$ equals the set of decision paths in $(T, \supseteq)$

containing x. By part (3) of lemma Appendix A.1, this set is equal to $W(x)$. By assumption (1), we obtain a map $f\colon V \to W$ satisfying, for any $T \in \mathcal{F}$ and $v \in V_T$, $f(v) = f_T(v)$. f is injective by lemma Appendix A.1, part (2), since $\mathcal{F}$ is a partition. f is also surjective, hence bijective, since every maximal chain $w \in W$ contains an element x for that there is $T \in \mathcal{F}$ with $x \in T$. Hence $w \in W(x)$ which is a subset of $W(V_T)$ by lemma Appendix A.1, part (3). Hence, there is $v \in V_T$ with $f(v) = f_T(v) = w$.

Let $x \in F$. Then there is $T \in \mathcal{F}$ with $x \in T$. Then $V_T \supseteq x$. Hence, $(\mathcal{P}f)(x) = (\mathcal{P}f_T)(x)$. By definition of f_T and part (3) of lemma Appendix A.1, we have

$$(\mathcal{P}f_T)(x) = \{w \in W(V_T) \mid x \in w\}.$$

Lemma Appendix A.1, part (3), implies that the latter set equals $W(x)$. Thus, $(\mathcal{P}f)(x) = W(x)$. □

Appendix A.2. Section 2

Proof of lemma 2.2. Let $\bar{\mathbf{X}} = \mathbf{X} \setminus \{\emptyset\}$, that is the set of all random moves $\mathbf{x}$ with $D_{\mathbf{x}} \neq \emptyset$. Then it is evident that $\bar{\mathbf{X}}$ satisfies properties 3a, 3b, 3c and 3d of definition 2.1, and is refined by $\mathbf{X}$. Hence, by the maximality property 3e in definition 2.1, $\bar{\mathbf{X}} = \mathbf{X}$. □

Proof of proposition 2.3. (Ad "order embedding"): Let $(\mathbf{y}_1, \omega_1), (\mathbf{y}_2, \omega_2) \in \mathbf{T} \bullet \Omega$. If both $\mathbf{y}_1, \mathbf{y}_2$ are elements of $\mathbf{X}$, the implication "$\Rightarrow$" is a consequence of the definition of $\geq_{\mathbf{X}}$. The implication "$\Leftarrow$" can be shown as follows: If $\mathbf{y}_1(\omega_1) \supseteq \mathbf{y}_2(\omega_2)$, then both moves belong to the same connected component of the rooted forest $(F, \supseteq)$ (whose existence axiom 1 ensures). By axioms 2 and 3a, $\omega_1 = \omega_2$. Then, axiom 3c implies that $\mathbf{y}_1 \geq_{\mathbf{X}} \mathbf{y}_2$, whence $\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_2$.

Else, for some $k = 1, 2$, $\mathbf{y}_k$ is a random terminal node so that $D_{\mathbf{y}_k} = \{\omega_k\}$ and $\mathbf{y}_k(\omega_k) = \{w_k\}$ for some $w_k \in W$. If $k = 1$, then first, $\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_2$ is equivalent to $\mathbf{y}_1 = \mathbf{y}_2$. Second, $\mathbf{y}_1(\omega_1) \supseteq \mathbf{y}_2(\omega_2)$ is equivalent to $\mathbf{y}_1(\omega_1) = \mathbf{y}_2(\omega_2)$, because $\mathbf{y}_1(\omega)$ is terminal in F. But then $\mathbf{y}_2$ must be a random terminal node as well, whence $\mathbf{y}_1 = \mathbf{y}_2$ and $\omega_1 = \omega_2$. Conversely, if these two equalities hold true, then $\mathbf{y}_1(\omega_1) = \mathbf{y}_2(\omega_2)$ follows.

If $k = 2$, then $\mathbf{y}_1(\omega_1) \supseteq \mathbf{y}_2(\omega_2)$ is equivalent to the conjunction of these three statements: $\omega_1 = \omega_2$, $\omega_2 \in D_{\mathbf{y}_1}$, and $w_k \in \mathbf{y}_1(\omega_1)$. By definition of $\geq_{\mathbf{T}}$, this is equivalent to $\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_2$ and $\omega_1 = \omega_2$.

(Ad "bijection"): Let $(\mathbf{y}_1, \omega_1), (\mathbf{y}_2, \omega_2) \in \mathbf{T} \bullet \Omega$ such that $\mathbf{y}_1(\omega_1) = \mathbf{y}_2(\omega_2)$. As the evaluation map is an order embedding, which we have proven just before, we obtain $\mathbf{y}_1 = \mathbf{y}_2$ and $\omega_1 = \omega_2$. Regarding surjectivity, let $y \in F$. If y is terminal, then there is $w \in W$ with $y = \{w\}$ because F is a decision forest over W.[16] Hence, $y = \mathbf{y}(\omega)$ for $\omega = \pi(y)$ and the random terminal node $\mathbf{y} = (\omega, w)$. If y is a move, then, by the definition of $\mathbf{X}$, property 3b, there is $\mathbf{x} \in \mathbf{X}$ and $\omega \in D_{\mathbf{x}}$ such that $\mathbf{x}(\omega) = y$. □

Proof of theorem 2.4. (Ad "poset"): It follows directly from the definitions that $\geq_{\mathbf{T}}$ defines a partial order on $\mathbf{T}$.

(Ad "forest"): Let $\mathbf{y}_0 \in \mathbf{T}$ and $\mathbf{y}_1, \mathbf{y}_2 \in \uparrow \mathbf{y}_0$. As $D_{\mathbf{y}_0} \neq \emptyset$ by lemma 2.2 and the definition of random terminal nodes, there is $\omega \in D_{\mathbf{y}_0}$, and we have $\omega \in D_{\mathbf{y}_k}$ and $\mathbf{y}_k(\omega) \supseteq \mathbf{y}_0(\omega)$ for both $k = 1, 2$. As $(F, \supseteq)$ is a forest, $\uparrow \mathbf{y}_0(\omega)$ is a chain, thus $\mathbf{y}_1(\omega) \supseteq \mathbf{y}_2(\omega)$ or $\mathbf{y}_2(\omega) \supseteq \mathbf{y}_1(\omega)$. Hence, by proposition 2.3, we get $\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_2$ or $\mathbf{y}_2 \geq_{\mathbf{T}} \mathbf{y}_1$. We conclude that $(\mathbf{T}, \geq_{\mathbf{T}})$ is a forest.

[16] Although this is discussed in [3], we give an argument here for the reader's convenience. Let f be the map from definition 1.3. If y is terminal, then $(\mathcal{P}f)(y) = W(y)$ contains exactly one element, namely $\uparrow y$. As f is injective, y must be a singleton.

(Ad "decision forest"): Let $\mathbf{y}_1, \mathbf{y}_2 \in \mathbf{T}$ be such that $\mathbf{y}_1 \neq \mathbf{y}_2$. If they are not comparable by $\geq_{\mathbf{T}}$, that is neither $\mathbf{y}_1 >_{\mathbf{T}} \mathbf{y}_2$ nor $\mathbf{y}_2 >_{\mathbf{T}} \mathbf{y}_1$, then any maximal chain $\mathbf{w}$ in $(\mathbf{T}, \geq_{\mathbf{T}})$ with $\mathbf{y}_1 \in \mathbf{w}$ satisfies $\mathbf{y}_2 \notin \mathbf{w}$. Using the axiom of choice in the form of the Hausdorff maximal principle, there exists indeed such a $\mathbf{w}$.

For symmetry reasons, it remains to consider the case in that we have $\mathbf{y}_1 >_{\mathbf{T}} \mathbf{y}_2$. Then, by proposition 2.3, we have for all $\omega' \in D_{\mathbf{y}_2}$, $\mathbf{y}_1(\omega') \supsetneq \mathbf{y}_2(\omega')$. As $D_{\mathbf{y}_2} \neq \emptyset$, we can select $\omega \in D_{\mathbf{y}_2}$. We can also select $w \in \mathbf{y}_1(\omega) \setminus \mathbf{y}_2(\omega)$. In particular, $\mathbf{y}_2(\omega) \notin \uparrow \{w\}$. As the latter set is a maximal chain in $(F, \supseteq)$, there is $y_3 \in \uparrow \{w\}$ such y_3 and $\mathbf{y}_2(\omega)$ cannot be compared by $\supseteq$. As $\mathbf{y}_1(\omega)$ is also an element of the chain $\uparrow \{w\}$, this implies $\mathbf{y}_1(\omega) \supsetneq y_3$. By proposition 2.3, there is $\mathbf{y}_3 \in \mathbf{T}$ such that $\omega \in D_{\mathbf{y}_3}$ and $y_3 = \mathbf{y}_3(\omega)$. This lemma also implies that $\mathbf{y}_2$ and $\mathbf{y}_3$ are not comparable via $\geq_{\mathbf{T}}$. Hence, as we have proven just above, there is a maximal chain $\mathbf{w}$ in $(\mathbf{T}, \geq_{\mathbf{T}})$ with $\mathbf{y}_2 \notin \mathbf{w}$ and $\mathbf{y}_3 \in \mathbf{w}$. By proposition 2.3, we obtain $\mathbf{y}_1 \geq_{\mathbf{T}} \mathbf{y}_3$, and using the fact that $(\mathbf{T}, \geq_{\mathbf{T}})$ is a forest, as already proven, we infer that $\mathbf{y}_1 \in \mathbf{w}$. Hence there is a maximal chain in $(\mathbf{T}, \geq_{\mathbf{T}})$ separating $\mathbf{y}_1$ and $\mathbf{y}_2$.

(Ad "rooted"): Suppose that $(\mathbf{T}, \geq_{\mathbf{T}})$ were not rooted. As $(F, \supseteq)$ is rooted and all roots are moves, by axioms 3b and 3d there is at least one random move whose image consists of roots. If $(\mathbf{T}, \geq_{\mathbf{T}})$ were not rooted, there would have to be at least two of them. Hence, there would be $\mathbf{x}_0, \mathbf{x}_1 \in \mathbf{X}$ whose images are roots of F and such that $D_{\mathbf{x}_0} \cap D_{\mathbf{x}_1} = \emptyset$, by axiom 3c. Then let $\bar{\mathbf{x}} = \mathbf{x}_0 \cup \mathbf{x}_1$,[17] and $\bar{\mathbf{X}} = (\mathbf{X} \setminus \{\mathbf{x}_0, \mathbf{x}_1\}) \cup \{\bar{\mathbf{x}}\}$.

Clearly, $D_{\bar{\mathbf{x}}} = D_{\mathbf{x}_0} \cup D_{\mathbf{x}_1}$ would be an element of $\mathscr{A}$ and $\bar{\mathbf{x}}$ would be a section of π on $D_{\bar{\mathbf{x}}}$. It is evident that $\bar{\mathbf{X}}$ would induce a covering of X and identify roots in the sense of axiom 3d. To show that $\bar{\mathbf{X}}$ would satisfy axiom 3c, it suffices to show that any $\mathbf{x} \in \mathbf{X} \setminus \{\mathbf{x}_0, \mathbf{x}_1\}$ admitting $\omega \in D_{\mathbf{x}} \cap D_{\bar{\mathbf{x}}}$ with $\bar{\mathbf{x}}(\omega) \supseteq \mathbf{x}(\omega)$ would have to satisfy $\bar{\mathbf{x}} \geq_{\bar{\mathbf{X}}} \mathbf{x}$. But this follows easily because the hypothesis on $\mathbf{x}$ and ω implies that there is $k = 0, 1$ such that $\omega \in D_{\mathbf{x}} \cap D_{\mathbf{x}_k}$ and $\mathbf{x}_k(\omega) \supseteq \mathbf{x}(\omega)$. Hence, $\mathbf{x}_k \geq_{\mathbf{X}} \mathbf{x}$ which directly yields $\bar{\mathbf{x}} \geq_{\bar{\mathbf{X}}} \mathbf{x}$. By construction, $\mathbf{X}$ would refine $\bar{\mathbf{X}}$. Thus, axiom 3e would imply the false statement $\bar{\mathbf{X}} = \mathbf{X}$.

We conclude that $(\mathbf{T}, \geq_{\mathbf{T}})$ must be rooted.

(Ad $\mathbf{X}$ = set of moves): Let $\mathbf{x} \in \mathbf{X}$. Then there is $\omega \in D_{\mathbf{x}}$ and, by axiom 3a, $\mathbf{x}(\omega) \in X$. Hence, there is $y \in F$ with $\mathbf{x}(\omega) \supsetneq y$. By proposition 2.3, there is unique $\mathbf{y} \in \mathbf{T}$ with $\omega \in D_{\mathbf{y}}$ and $y = \mathbf{y}(\omega)$, and whence, by the same proposition, $\mathbf{x} >_{\mathbf{T}} \mathbf{y}$. Hence, $\mathbf{x}$ is a move in $(\mathbf{T}, \geq_{\mathbf{T}})$.

On the other hand, the elements of $\mathbf{T} \setminus \mathbf{X}$, i.e., the random terminal nodes, are terminal nodes in $(\mathbf{T}, \geq_{\mathbf{T}})$ directly by definition of $\geq_{\mathbf{T}}$. Hence, $\mathbf{X}$ is the set of moves in $(\mathbf{T}, \geq_{\mathbf{T}})$. □

We continue with the verification of the presented examples of stochastic decision forests.

Proof of lemma 2.6. $(\Omega, \mathscr{A})$ is a measurable space. F is clearly a W-poset and the map assigning to any $w \in W$ the chain $\uparrow \{w\}$ is easily seen to be a bijection between W and the set of maximal chains in $(F, \supseteq)$. Moreover, any chain in $(F, \supseteq)$ contains a maximal element, hence, by [3, theorem 3], expressed in the language of the present paper, F a decision forest over W (axiom 1).

The connected components of $(F, \supseteq)$ are easily identified as

$$\{\mathbf{x}_k(\omega) \mid k = 0, 1, 2\} \cup \{\{(\omega, k, m)\} \mid k, m \in \{1, 2\}\},$$

ranging over $\omega \in \Omega$. π is well-defined because all nodes are non-empty subsets of W and its definition does not depend on the choice of the element whose first entry is evaluated. The connected component spelled out above equals $\pi^{-1}(\{\omega\})$ indeed. Hence, axiom 2 is satisfied.

[17] That is, $\bar{\mathbf{x}} \colon D_{\mathbf{x}_0} \cup D_{\mathbf{x}_1} \to X$ such that $\bar{\mathbf{x}}(\omega) = \mathbf{x}_k(\omega)$ for $\omega \in D_{\mathbf{x}_k}$, for both $k = 0, 1$.

For all $k = 0, 1, 2$, $\mathbf{x}_k$ is defined on Ω which is an event. Clearly, $\mathbf{x}_k$ is a section of π. The union of the images of $\mathbf{X}$'s elements is the set of nodes with at least two elements. As $\{\{w\} \mid w \in W\} \subseteq F$, this union equals X. Thus axioms 3a and 3b are satisfied.

Furthermore, the relations between the images of all presumed random moves, put aside equalities, amount to

$$\forall \omega \in \Omega \forall k = 1, 2\colon \quad \mathbf{x}_0(\omega) \supseteq \mathbf{x}_k(\omega).$$

Hence, put aside equalities, $\geq_{\mathbf{X}}$ is given by $\mathbf{x}_0 \geq_{\mathbf{X}} \mathbf{x}_k$, for $k = 1, 2$. Moreover, the roots of F are given by the image of $\mathbf{x}_0$. It follows that $\mathbf{X}$ satisfies axioms 3c and 3d.

As the domains of all $\mathbf{x} \in \mathbf{X}$ are equal to the sample space Ω, the maximality axiom 3e is satisfied. As $\mathbf{X}$ is finite, the separability condition 3f is fulfilled. $\square$

Proof of lemma 2.7. The proof is almost completely analogous to the preceding one, up to small modifications. We only comment on those. Thus, we omit the first two axioms.

For all $k = 0, 1$, $\mathbf{x}_k$ is defined on Ω as in the previous lemma. For $k = 2$, it has to be noted that $\{\omega_2\} \in \mathscr{A}$. Clearly, $\mathbf{x}'$ is a section of π for all $\mathbf{x}' \in \mathbf{X}'$. The union of the images of the elements of $\mathbf{X}'$ is the set of nodes with at least two elements. Again, as $\{\{w'\} \mid w' \in W\} \subseteq F'$, this union equals X'. It follows that $\mathbf{X}'$ satisfies axioms 3a and 3b.

Furthermore, the relations between the images of all presumed random moves, put aside equalities, amount to

$$\forall \omega \in \Omega\colon \quad \mathbf{x}'_0(\omega) \supseteq \mathbf{x}'_1(\omega), \qquad \mathbf{x}'_0(\omega_2) \supseteq \mathbf{x}'_2(\omega_2).$$

Hence, put aside equalities, $\geq_{\mathbf{X}'}$ is given by $\mathbf{x}'_0 \geq_{\mathbf{X}'} \mathbf{x}'_1$ and $\mathbf{x}'_0 \geq_{\mathbf{X}'} \mathbf{x}'_2$. The roots of F' are given by the image of $\mathbf{x}'_0$. It follows that $\mathbf{X}'$ satisfied axioms 3c and 3d.

The maximality axiom 3e requires only a little bit more justification this time. If $\bar{\mathbf{X}}'$ satisfies the first four axioms 3a, 3b, 3c, and 3d, and $\mathbf{X}'$ refines $\bar{\mathbf{X}}'$, then $\mathbf{x}'_0, \mathbf{x}'_1 \in \bar{\mathbf{X}}'$ because of the full domain of these presumed random moves and axioms 3a and 3b. By proposition 2.3 and remark 2.5, there there must be one and only one other element of $\bar{\mathbf{X}}'$, namely $\mathbf{x}'_2$. Hence, $\bar{\mathbf{X}}' = \mathbf{X}'$. The separability condition 3f is again fulfilled because of the finiteness of the situation. $\square$

Proof of lemma 2.8. (Ad "$\Rightarrow$"): Suppose AP.W1 to hold and let $x \in F$ be not a singleton. Then, there are $(t, w) \in \mathbb{T} \times W$ with $x = x_t(w)$. Let $(t', w') \in \mathbb{T} \times W$ such that $x = x_{t'}(w')$. Hence, $w \in x_{t'}(w')$ and thus $x = x_{t'}(w)$. By assumption AP.W1, $t = t'$. Hence, $\mathbb{T}_x = \{t\}$.

(Ad "$\Leftarrow$") Suppose the right-hand criterion to be satisfied. Let $w \in W$ and $t, u \in \mathbb{T}$ with $t \neq u$ and $x_t(w) = x_u(w)$. Hence, $\mathbb{T}_{x_t(w)}$ is not a singleton, and by our hypothesis $x_t(w)$ must be a singleton. As $w \in x_t(w)$, we get $x_t(w) = \{w\}$. $\square$

Proof of lemma 2.9. Let $x \in F$ such that $\mathbb{T}_x = \{t\}$ for some non-maximal $t \in \mathbb{T}$, and let $w \in x$. Then there is $u \in \mathbb{T}$ with $t < u$, whence $x = x_t(w) \supsetneq x_u(w)$. As $w \in x_u(w)$, x contains at least two elements. $\square$

Proof of theorem 2.11. (Ad axiom 1): As $F \subseteq \mathcal{P}(W)$, F is a W-poset. We show that it is a "bounded" and "irreducible" "W-set tree", in the language of [3], and that any principal up-set in $(F, \supseteq)$ has a maximal element. From this, using [3, theorem 3], we obtain, expressed in the language of the present paper, that F is decision forest over W.

A "W-set tree" is a W-poset satisfying "trivial intersection" and "separability" (see [3, definition 3]). Regarding "trivial intersection" (defined in [3, (4), p. 768]), let $x, y \in F$ be such that $x \cap y \neq \emptyset$. Then there is $w \in W$ such that $w \in x \cap y$. Hence, x and y are, respectively, of the form $\{w\}$ or $x_t(w)$, for some $t \in \mathbb{T}$. In each of the four possible cases, we get $x \supseteq y$ or $y \supseteq x$. Hence, the "trivial intersection" property is satisfied. Regarding "separability" (defined in [3, (8), p. 773]), let $x, y \in F$

such that $x \supsetneq y$. Then there is $w \in y$, whence $w \in x$. x has at least two elements, thus $x = x_t(w)$ for some $t \in \mathbb{T}$. Let $w' \in x \setminus y$. Then $x \supseteq \{w'\}$ and $y \cap \{w'\} = \emptyset$. Hence, the "separability" property is satisfied, and F is a W-set tree.

Regarding "irreducibility", let $w_0, w_1 \in W$ with $w_0 \neq w_1$. Then $\{w_k\} \in F$ satisfy $w_k \in \{w_k\} \setminus \{w_{1-k}\}$ for both $k = 0, 1$. Whence irreducibility.

Regarding "boundedness", let c be a chain in $(F, \supseteq)$. It suffices to consider the case where $c \neq \emptyset$. First, consider the case that there is $w \in W$ with $\{w\} \in c$. In that case, any $x \in c$ is non-empty and hence the chain property of c implies $w \in x$. Second, if there is no $w \in W$ with $\{w\} \in c$, then there must be a unique $\omega \in \Omega$ such that

$$\mathbb{T}' = \{t' \in \mathbb{T} \mid \exists f_{t'} \in \mathbb{A}^{\mathbb{T}} \colon x_{t'}(\omega, f_{t'}) \in c\}$$

is non-empty. Uniqueness holds true because if $t_1, t_2 \in \mathbb{T}$ and $(\omega_1, f_1), (\omega_2, f_2) \in W$ satisfy $x_{t_1}(\omega_1, f_1) \supseteq x_{t_2}(\omega_2, f_2)$, then, by definition, $\omega_1 = \omega_2$.

Let $\sup \mathbb{T}'$ denote the supremum of $\mathbb{T}'$ in $\bar{\mathbb{R}}_+ = [0, +\infty]$. As c is a chain and by definition of $\mathbb{T}'$, for all $t \in \mathbb{T}$ with $t < \sup \mathbb{T}'$, there is a unique $a(t) \in \mathbb{A}$ such that for all $t' \in \mathbb{T}'$ with $t < t'$ and $f_{t'} \in \mathbb{A}^{\mathbb{T}}$ satisfying $x_{t'}(\omega, f_{t'}) \in c$ we have $f_{t'}(t) = a(t)$. Let $\tilde{f} \in \mathbb{A}^{\mathbb{T}}$ be such that $\tilde{f}(t) = a(t)$ for all $t \in [0, \sup \mathbb{T}') \cap \mathbb{T}$. Hence, all $t' \in \mathbb{T}'$ satisfy $x_{t'}(\omega, \tilde{f}) \in c \subseteq F$. By assumption AP.W2, there is $f \in \mathbb{A}^{\mathbb{T}}$ such that $(\omega, f) \in W$ and $x_{t'}(\omega, f) = x_{t'}(\omega, \tilde{f}) \in c$ for all $t' \in \mathbb{T}'$.

Hence, as c contains no singleton, we get

$$(*) \qquad c = \{x_{t'}(\omega, f) \mid t' \in \mathbb{T}'\}.$$

Hence, for all $x \in c$, we have $(\omega, f) \in x$. The proof of "boundedness" is complete.

It remains to be shown that any principal up-set contains a maximal element. The crucial point here is that for any $x \in F$, there is $\omega \in \Omega$ such that $x \subseteq (\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W$. Moreover, for any $w \in x$, $(\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W = x_0(w)$. Thus $x_0(w) \in \uparrow x$, and if $y \in \uparrow x_0(w)$, then $y = x_0(w)$. Thus, $x_0(w) \in \uparrow x$ is a maximal element for $(F, \supseteq)$.

(Ad axiom 2): Recall from the preceding argument that any $x \in F$ is contained in $(\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W$ for some $\omega \in \Omega$. Hence, $\pi(x)$ is well-defined and equal to this uniquely determined ω. Let $x, y \in F$. Then $\pi(x) = \pi(y)$ iff there is $\omega \in \Omega$ such that $x, y \subseteq (\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W$. As $(\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W$ is a root of $(F, \supseteq)$, as seen just above, this is equivalent to the statement that x, y belong to the same connected component.

(Ad axiom 3a): By definition of the sets $D_{t,f}$, $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$, and in view of the fact that F contains all singletons in $\mathcal{P}(W)$, all elements of $\mathbf{X}$ map into the set X of moves of the decision forest F over W. Moreover, by assumption AP.W0, $D_{t,f} \in \mathscr{A}$. Furthermore, any $\mathbf{x} \in \mathbf{X}$ is a section of π. For this represent $\mathbf{x}$ as $\mathbf{x}_t(f)$ for $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$ such that $D_{t,f} \neq \emptyset$, and let $\omega \in D_{t,f}$. Then,

$$\pi(\mathbf{x}_t(f)(\omega)) = \pi(x_t(\omega, f)) = \omega,$$

because ω is the first component of $(\omega, f) \in x_t(\omega, f)$.

(Ad axiom 3b): Let $x \in X$. Then x has at least two elements, because F is a decision forest over W. Hence, there are $(\omega, f) \in W$ and $t \in \mathbb{T}$ such that $x = x_t(\omega, f)$. Hence, $\omega \in D_{t,f}$, and $x = \mathbf{x}_t(f)(\omega)$.

(Ad axiom 3c): Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}$ such that there is $\omega \in D_{\mathbf{x}_1} \cap D_{\mathbf{x}_2}$ with $\mathbf{x}_1(\omega) \supseteq \mathbf{x}_2(\omega)$. Hence, there are $t_1, t_2 \in \mathbb{T}$ and $f \in \mathbb{A}^{\mathbb{T}}$ such that $\omega \in D_{t_k,f}$ and $\mathbf{x}_k = \mathbf{x}_{t_k}(\omega, f)$, for both $k = 1, 2$. If we had $t_1 > t_2$, then, by construction of the nodes and by assumption AP.W1, $\mathbf{x}_1(\omega) = x_{t_1}(\omega, f) \subsetneq x_{t_2}(\omega, f) = \mathbf{x}_2(\omega)$, because they contain at least two elements as moves – a contradiction. Hence, $t_1 \leq t_2$.

Let $\omega' \in D_{t_2,f}$. By definition, $x_{t_2}(\omega', f)$ is a move. As $t_1 \leq t_2$, $x_{t_1}(\omega', f) \supseteq x_{t_2}(\omega', f)$, whence $\omega' \in D_{t_1,f}$. In particular, $\mathbf{x}_1(f)(\omega') \supseteq \mathbf{x}_2(f)(\omega')$. Thus, $\mathbf{x}_1 \geq_{\mathbf{X}} \mathbf{x}_2$.

(Ad axiom 3d): Let $\mathbf{x} \in \mathbf{X}$ admitting $\omega \in D_{\mathbf{x}}$ with $\mathbf{x}(\omega) = (\{\omega\} \times \mathbb{A}^{\mathbb{T}}) \cap W$. Hence, for any $f \in \mathbb{A}^{\mathbb{T}}$, $D_{0,f} \neq \emptyset$ and $\mathbf{x}_0(f)(\omega) = \mathbf{x}(\omega)$. Let us pick an $f \in \mathbb{A}^{\mathbb{T}}$. Hence, $\mathbf{x}_0(f) \in \mathbf{X}$ and, by axiom 3c, $\mathbf{x} = \mathbf{x}_0(f)$. Thus, for all $\omega' \in D_{\mathbf{x}} = D_{0,f}$, we have $\mathbf{x}(\omega') = x_0(\omega', f) = (\{\omega'\} \times \mathbb{A}^{\mathbb{T}}) \cap W$.

(Ad axiom 3e): Let $\bar{\mathbf{X}}$ be a set satisfying the first four axioms of 3 that is refined by $\mathbf{X}$. By definition of the latter, for any $\bar{\mathbf{x}} \in \bar{\mathbf{X}}$ there is $P_{\bar{\mathbf{x}}} \subseteq \mathbf{X}$ such that $\bar{\mathbf{x}} = \bigcup P_{\bar{\mathbf{x}}}$. As $\bar{\mathbf{X}}$ and $\mathbf{X}$ induce partitions of X, by proposition 2.3 and remark 2.5, and as $D_{\mathbf{x}} \neq \emptyset$ for all $\mathbf{x} \in \mathbf{X}$ by construction of $\mathbf{X}$, we infer that there is a unique map $b\colon \mathbf{X} \to \bar{\mathbf{X}}, \mathbf{x} \mapsto b(\mathbf{x})$ such that $D_{\mathbf{x}} \subseteq D_{b(\mathbf{x})}$ and $b(\mathbf{x})|_{D_{\mathbf{x}}} = \mathbf{x}$, for all $\mathbf{x} \in \mathbf{X}$. For each pair $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$ with $D_{t,f} \neq \emptyset$, denote $\bar{\mathbf{x}}_t(f) = b(\mathbf{x}_t(f))$ the image of $\mathbf{x}_t(f)$ under this map.

Our aim is to show that for all $\bar{\mathbf{x}} \in \bar{\mathbf{X}}$, we have $P_{\bar{\mathbf{x}}} = \{\bar{\mathbf{x}}\}$. If this were not the case, there would be $\bar{\mathbf{x}} \in \bar{\mathbf{X}}$ such that $P_{\bar{\mathbf{x}}}$ has at least two elements. Hence, there would be $t, u \in \mathbb{T}$ and $f, g \in \mathbb{A}^{\mathbb{T}}$ with $D_{t,f}, D_{u,g} \neq \emptyset$, $\mathbf{x}_t(f) \neq \mathbf{x}_u(g)$, and $\mathbf{x}_t(f), \mathbf{x}_u(g) \in P_{\bar{\mathbf{x}}}$. This would imply

$$(\dagger) \qquad D_{t,f} \cap D_{u,g} = \emptyset.$$

Indeed, otherwise there would be $\omega \in D_{t,f} \cap D_{u,g}$ for that that definition of $\bar{x}_t(f)$ combined with $(\bullet)$ would imply $x_t(\omega, f) = x_u(\omega, g)$. Hence, for $v = t \wedge u$, we would have $f|_{[0,v)_{\mathbb{T}}} = g|_{[0,v)_{\mathbb{T}}}$ and there would be $w \in W$ with $x_t(\omega, f) = x_t(w) = x_u(w) = x_u(\omega, g)$. As these are moves, assumption AP.W1 would imply $t = u = v$, whence $\mathbf{x}_t(f) = \mathbf{x}_u(g)$ which is absurd. Hence, we would necessarily have $D_{t,f} \cap D_{u,g} = \emptyset$.

We can assume, without loss of generality, that $t \leq u$. There would be two cases. First, $D_{t,f} \cap D_{t,g} \neq \emptyset$. Then, $f|_{[0,t)_{\mathbb{T}}} \neq g|_{[0,t)_{\mathbb{T}}}$, because else $D_{t,f} = D_{t,g}$ whence $D_{u,g} \subseteq D_{t,g} = D_{t,f}$ which contradicts $(\dagger)$. The second case would be $D_{t,f} \cap D_{t,g} = \emptyset$. Then, by assumption AP.W3, there would be $v \in [0, t)_{\mathbb{T}}$ such that

$$(\circ) \qquad D_{v,f} \cap D_{v,g} \neq \emptyset, \quad \text{and} \quad f|_{[0,v)_{\mathbb{T}}} \neq g|_{[0,v)_{\mathbb{T}}}.$$

We conclude that in both cases, there would be $v \in [0, t]_{\mathbb{T}}$ satisfying $(\circ)$.

Let $\omega' \in D_{v,f} \cap D_{v,g}$. Then, $x_v(\omega', f)$ and $x_v(\omega', g)$ would not be comparable in $(F, \supseteq)$ since equality would imply $f|_{[0,v)_{\mathbb{T}}} = g|_{[0,v)_{\mathbb{T}}}$. Hence, by the definition of $\geq_{\bar{\mathbf{X}}}$, the same would be true for $\bar{\mathbf{x}}_v(f)$ and $\bar{\mathbf{x}}_v(g)$ in $(\bar{\mathbf{X}}, \geq_{\bar{\mathbf{X}}})$. But, on the other hand, $\mathbf{x}_v(f) \geq_{\mathbf{X}} \mathbf{x}_t(f)$ and $\mathbf{x}_v(g) \geq_{\mathbf{X}} \mathbf{x}_u(g)$, whence by definition of $\geq_{\mathbf{X}}$ and axiom 3c applied to $\bar{\mathbf{X}}$, we get $\bar{\mathbf{x}}_v(f), \bar{\mathbf{x}}_v(g) \in \uparrow \bar{\mathbf{x}}_t(f)$. As $(\bar{\mathbf{X}}, \geq_{\bar{\mathbf{X}}})$ is a tree by proposition 2.3 and remark 2.5, $\bar{\mathbf{x}}_v(f)$ and $\bar{\mathbf{x}}_v(g)$ would have to be comparable, a contradiction. Hence, the initial assumption was false, and we infer that $P_{\bar{\mathbf{x}}} = \{\bar{\mathbf{x}}\}$ for all $\bar{\mathbf{x}} \in \bar{\mathbf{X}}$.

This in turn implies that $\mathbf{x} = b(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{X}$. Hence, we have both $\bar{\mathbf{X}} \subseteq \mathbf{X}$ and $\mathbf{X} \subseteq \bar{\mathbf{X}}$, whence $\mathbf{X} = \bar{\mathbf{X}}$.

(Ad axiom 3f): As $\mathbb{T}$ is embedded in $\mathbb{R}_+$, there is countable $\mathbb{T}_0 \subseteq \mathbb{T}$, dense in $\mathbb{T}$ with respect to the induced topology on $\mathbb{T}$. Let $\mathbf{x} \in \mathbf{X}$. Then, by construction of the nodes there are $t \in \mathbb{T}$ and $f \in \mathbb{A}^{\mathbb{T}}$ such that $\mathbf{x} = \mathbf{x}_t(f)$. Let

$$\mathbf{X}_0 = \{\mathbf{x}_{t_0}(f) \mid t_0 \in \mathbb{T}_0,\ t_0 < t\}.$$

Let $\mathbf{x}' \in \mathbf{X}$ such that $\mathbf{x}' <_{\mathbf{X}} \mathbf{x}$. Then, by construction of the nodes and lemma 2.13, there is $t' \in \mathbb{T}$ with $t' < t$ such that $\mathbf{x}' = \mathbf{x}_{t'}(f)$. Hence, there is $t_0 \in \mathbb{T}_0$ with $t' \leq t_0 \leq t$, and we easily obtain, directly applying the definitions, $D_{t',f} \supseteq D_{t_0,f} \supseteq D_{t,f}$ and for all $\omega \in D_{t,f}$, $x_{t_0}(\omega, f) \supseteq x_t(\omega, f)$, and for all $\omega_0 \in D_{t_0,f}$, $x_{t'}(\omega_0, f) \supseteq x_{t_0}(\omega_0, f)$. Thus $D_{t_0,f} \neq \emptyset$, $\mathbf{x}_{t_0}(f) \in \mathbf{X}$, and $\mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}_{t_0}(f) \geq_{\mathbf{X}} \mathbf{x}$. $\square$

Proof of lemma 2.13. Let $x, y \in X$ with $x \supsetneq y$. Then, by definition of $\mathfrak{t}$ and the construction of nodes in action path SDF, there is $w \in W$ such that $x = x_{\mathfrak{t}(x)}(w)$ and $y = x_{\mathfrak{t}(y)}(w)$. Indeed, y is non-empty, and any $w \in y$ satisfies the preceding equalities. If we had $\mathfrak{t}(x) \geq \mathfrak{t}(y)$, then $y = x_{\mathfrak{t}(y)}(w) \supseteq x_{\mathfrak{t}(x)}(w) = x$, in contradiction to the choice of x and y. Hence, by totality of the order on $\mathbb{T}$, $\mathfrak{t}(x) < \mathfrak{t}(y)$.

Let $\mathbf{x} \in \mathbf{X}$. Then there are $t \in \mathbb{T}$ and $f \in \mathbb{A}^{\mathbb{T}}$ such that $D_{t,f} \neq \emptyset$ with $\mathbf{x} = \mathbf{x}_t(f)$. Hence, for all $\omega \in D_{\mathbf{x}} = D_{t,f}$:

$$\mathfrak{t}(\mathbf{x}(\omega)) = \mathfrak{t}(x_t(\omega, f)) = t,$$

which does not depend on ω. □

Proof of lemma 2.14. Take $\mathbb{A} = \{1, 2\}$ and $\mathbb{T} = \{0, 1\}$. Recall that $\Omega = \{\omega_1, \omega_2\}$, with $\omega_k = k$ for $k \in \{1, 2\}$, and $\mathscr{A} = \mathcal{P}(\Omega)$.

For the first example, we may set $W = \Omega \times \mathbb{A}^{\mathbb{T}}$. Any $(\omega, f) \in W$ can be identified with the triple $(\omega, f(0), f(1))$, so that W corresponds to the set denoted by W in the example in subsection 2.3. Under this identification, it is easily verified that $D_{t,f} = \Omega$ for all $f \in \mathbb{A}^{\mathbb{T}}$ and $t \in \mathbb{T}$, $\mathbf{x}_0 = \mathbf{x}_0(f)$ for all $f \in \mathbb{A}^{\mathbb{T}}$, and $\mathbf{x}_k = \mathbf{x}_1(f)$ for all $f \in \mathbb{A}^{\mathbb{T}}$ with $f(0) = k$, for both $k = 1, 2$. Assumptions AP.W0 to AP.W3 are readily verified. Under this identification, the action path construction yields exactly the objects F, π, and $\mathbf{X}$ from the example in subsection 2.3.

For the second example, let W' as in subsection 2.3. Interpret every triple $(\omega, k, m) \in W'$ as the pair (ω, f) where $f \in \mathbb{A}^{\mathbb{T}}$ is given by $f(0) = k$ and $f(1) = m$. Interpret the pair $(\omega_1, 2)$ as the map $f \colon \mathbb{T} \to \{0, 1, 2\}$ with $f(0) = 2$ and $f(1) = 0$. Then, $W \subseteq \mathbb{A}'^{\mathbb{T}}$ with $\mathbb{A}' = \mathbb{A} \cup \{0\}$, where 0 is a placeholder for inaction. Furthermore, $D_{t,f} = \Omega$ for all $(t, f) \in \mathbb{T} \times \mathbb{A}'^{\mathbb{T}}$ with $t = 0$, or $t = 1$ and $f(0) = 1$; and $D_{1,f} = \{\omega_2\}$ for $f \colon \mathbb{T} \to \mathbb{A}'$ with $f(0) = 2$; for all other pairs $(t, f) \in \mathbb{T} \times \mathbb{A}'^{\mathbb{T}}$ we have $D_{t,f} = \emptyset$. Moreover, $\mathbf{x}'_0 = \mathbf{x}_0(f)$ for all $f \in \mathbb{A}'^{\mathbb{T}}$; and $\mathbf{x}'_1 = \mathbf{x}_1(f)$ for all $f \colon \mathbb{T} \to \mathbb{A}'$ with $f(0) = 1$; and $\mathbf{x}'_2 = \mathbf{x}_1(f)$ for $f \colon \mathbb{T} \to \mathbb{A}'$ with $f(0) = 2$. Assumptions AP.W0 to AP.W3 are readily verified. Under these identifications, the action path construction yields exactly the objects F', π', and $\mathbf{X}'$ from the example in subsection 2.3. □

Proof of the claims in example 2.15. (The case $W = \Omega \times \mathbb{A}^{\mathbb{T}}$): We show that for $W = \Omega \times \mathbb{A}^{\mathbb{T}}$, W induces an action path SDF. First, we note that for any $\omega \in \Omega$, there is $f \in \mathbb{A}^{\mathbb{T}}$ such that $(\omega, f) \in W$.

Clearly, for all $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$, $D_{t,f} = \Omega$, whence AP.W0.

Regarding AP.W1, let $w = (\omega, f) \in W$ and $t, u \in \mathbb{T}$ with $t \neq u$ such that $x_t(w) = x_u(w)$. Without loss of generality, assume $t < u$. If $\mathbb{A}$ is a singleton, then $\mathbb{A}^{\mathbb{T}}$ is a singleton as well. Hence, $x_t(w) = \{w\}$. But $\mathbb{A}$ must be a singleton, since else there would be $g \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, g) \in x_t(w)$ and $g(t) \neq f(t)$. Thus, $(\omega, g) \in x_t(w) \setminus x_u(w)$ – a contradiction.

Regarding AP.W2, let $\omega \in \Omega$, $\tilde{f} \in \mathbb{A}^{\mathbb{T}}$, and $\mathbb{T}' \subseteq \mathbb{T}$ satisfying $x_t(\omega, \tilde{f}) \in F$ for all $t \in \mathbb{T}'$. Then we already have $(\omega, \tilde{f}) \in W$.

AP.W3 is clearly satisfied because $D_{t,f} = \Omega$ for all $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$.

(The timing problem): Clearly, for any $\omega \in \Omega$ there is $f \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, f) \in W$, e.g., the constant map taking only the value 1 at all times, in all components.

Regarding assumption AP.W0, for all $(t, f) \in \mathbb{A}^{\mathbb{T}}$, $D_{t,f} = \emptyset$ if f is not decreasing on $[0, t)_{\mathbb{T}}$ or if $f(t-) = 0$. Else, $D_{t,f} = \Omega$. Hence, AP.W0 is satisfied.

Regarding assumption AP.W1, let $w = (\omega, f) \in W$ and $t, u \in \mathbb{T}$ with $t < u$ such that $x_t(w) = x_u(w)$. Hence, for all decreasing $g \colon \mathbb{T} \to \mathbb{A}$ with $g|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}}$ we must have $f(v) = g(v)$ for all $v \in [t, u)_{\mathbb{T}}$. Thus, we must have $f(t-) = 0$. Hence, $x_t(w) = \{(\omega, f)\}$.

Regarding assumption AP.W2, let $\omega \in \Omega$, $\tilde{f} \in \mathbb{A}^{\mathbb{T}}$, and $\mathbb{T}' \subseteq \mathbb{T}$ such that $x_t(\omega, \tilde{f}) \in F$ for all $t \in \mathbb{T}'$. Hence, $\tilde{f}$ is decreasing on $[0, \sup \mathbb{T}')_{\bar{\mathbb{R}}_+} \cap \mathbb{T}$ (where, in the context of the order on $\mathbb{T}$, we have $\sup\emptyset = 0$). Hence, there is decreasing $f\colon \mathbb{T} \to \mathbb{A}$ such that $f|_{[0,t)_{\mathbb{T}}} = \tilde{f}|_{[0,t)_{\mathbb{T}}}$ for all $t \in \mathbb{T}'$. By construction, $(\omega, f) \in W$.

Regarding assumption AP.W3, it is sufficient to note that $D_{t,f}$ equals $\emptyset$ or Ω, for all $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{T}}$ which has been established earlier.

(Ad American up-and-out option): Let $\mathbb{A} = \{0, 1\}$ and $\mathbb{T} = \mathbb{R}_+$. Clearly, for any $\omega \in \Omega$, there is $f\colon \mathbb{R}_+ \to \mathbb{A}$ with $(\omega, f) \in W$, namely, the constant path with value 1.

Regarding assumption AP.W0, let $(t, f) \in \mathbb{T} \times \mathbb{A}^{\mathbb{R}_+}$ and $\omega \in \Omega$. Then, we have $\omega \in D_{t,f}$ iff $x_t(\omega, f)$ has at least two elements. This is equivalent to the fact that f is decreasing on $[0, t)$, $f(t-) = 1$, and $P_u(\omega) < 2$ for all $u \in [0, t]$. Whence the claimed representation of $D_{t,f}$ and the fact that $D_{t,f} \in \mathscr{A}$, because P has continuous paths so that $\max_{u\in[0,t]} P_u$ is $\mathscr{A}$-measurable since P_u is so for all real (and in particular rational) $u \geq 0$.

Regarding assumption AP.W1, let $w = (\omega, f) \in W$ and $t, u \in \mathbb{R}_+$ with $t < u$ such that $x_t(w) = x_u(w)$. Hence, for all decreasing $g\colon \mathbb{R}_+ \to \{0, 1\}$ with $g|_{[0,t)} = f|_{[0,t)}$ and $(\omega, g) \in W$, we must have

$$(*) \qquad f(v) = g(v), \quad \text{for all} \;\; v \in [t, u).$$

If f is constant to 1, this implies the existence of $t_0 \in [0, t]$ such that $P_{t_0}(\omega) \geq 2$. If such t_0 did not exist, the map $g\colon \mathbb{R}_+ \to \{0, 1\}$ given by $g(t') = 1\{t' < t\}$ would violate condition $(*)$ above. Hence, $x_t(w) = \{w\}$. If f takes the value 0 and $t^*_f < t$, then clearly $x_t(w) = \{w\}$. The remaining case, namely that f takes the value 0 and $t^*_f \geq t$, cannot arise. Indeed, if it did, for both $a \in \{0, 1\}$, the map $g_a\colon \mathbb{R}_+ \to \{0, 1\}$ given by $g_a(t') = 1$ for $t' < t$, $g_a(t) = a$, and $g_a(t') = 0$ for $t' > t$ would satisfy $t^*_{g_a} = t$. Moreover, the fact that $(\omega, f) \in W$ would imply

$$\max_{t'\in[0,t]} P_{t'}(\omega) \leq \max_{t'\in[0,t^*_f]} P_{t'}(\omega) < 2,$$

hence $(\omega, g_a) \in W$. In particular, $(*)$ would imply that $a = g_a(t) = f(t)$, for both $a \in \{0, 1\}$ which is impossible. Hence, in any possible case we have $x_t(w) = \{w\}$, thus assumption AP.W1 is satisfied.

Regarding assumption AP.W2, let $\omega \in \Omega$, $\tilde{f} \in \mathbb{A}^{\mathbb{R}_+}$, and $\mathbb{T}' \subseteq \mathbb{R}_+$ such that $x_t(\omega, \tilde{f}) \in F$ for all $t \in \mathbb{T}'$. Hence, $\tilde{f}$ is decreasing on $[0, \sup\mathbb{T}')$ (where, in the context of the order on $\mathbb{R}_+$, we have $\sup\emptyset = 0$), and if $\tilde{f}|_{[0,\sup\mathbb{T}')}$ attains 0, then $P_u(\omega) < 2$ for all $u \in [0, t^*_{\tilde{f}}]$. Let $f\colon \mathbb{R}_+ \to \{0, 1\}$ be given by $f|_{[0,\sup\mathbb{T}')} = \tilde{f}|_{[0,\sup\mathbb{T}')}$ and, for all $u \in [\sup\mathbb{T}', \infty)$, $f(u) = f(\sup\mathbb{T}'-)$ if $\sup\mathbb{T}' > 0$ and $f(u) = 1$ if $\sup\mathbb{T}' = 0$. By construction, $f|_{[0,t)} = \tilde{f}|_{[0,t)}$ for all $t \in \mathbb{T}'$. Further, f is decreasing. Moreover, if f attains the value 0, then it does so on $[0, \sup\mathbb{T}')$ and so does $\tilde{f}$, and $t^*_f = t^*_{\tilde{f}}$. Hence, $(\omega, f) \in W$. We conclude that assumption AP.W2 is satisfied.

Regarding assumption AP.W3, let $t \in \mathbb{R}_+$ and $f, g \in \mathbb{A}^{\mathbb{T}}$ such that $D_{t,f}, D_{t,g} \neq \emptyset$. Then, as shown earlier in this proof, $f(t-) = g(t-) = 1$ and

$$D_{t,f} = \{\omega \in \Omega \mid \max_{u\in[0,t]} P_u(\omega) < 2\} = D_{t,g}.$$

Hence, $D_{t,f} \cap D_{t,g} \neq \emptyset$. Thus, assumption AP.W3 is trivially satisfied. □

Appendix A.3. Section 3

Proof of lemma 3.2. There are exactly two (σ-)algebras on Ω: the discrete and the trivial one, that is, $\mathcal{P}(\Omega)$ and $\{\Omega, \emptyset\}$. In the present situation, we have $D_{\mathbf{x}} = \Omega$ for all $\mathbf{x} \in \mathbf{X}$. Hence, by definition, the set of exogenous information structures $\mathscr{F}$ on $(F, \pi, \mathbf{X})$ is given by all families $\mathscr{F} = (\mathscr{F}_{\mathbf{x}})_{\mathbf{x}\in\mathbf{X}}$ of σ-algebras on Ω such that $\mathscr{F}_{\mathbf{x}_0} \subseteq \mathscr{F}_{\mathbf{x}_1} \cap \mathscr{F}_{\mathbf{x}_2}$. The claim follows easily from this. □

Proof of lemma 3.3. There is exactly one (σ-)algebra on $\{\omega_2\}$, namely $\{\{\omega_2\}, \emptyset\}$. On Ω, there are exactly two σ-algebras, namely $\mathcal{P}(\Omega)$ and $\{\Omega, \emptyset\}$, as in the preceding proof. Hence, by definition, the set of exogenous information structures $\mathscr{F}'$ on $(F', \pi', \mathbf{X}')$ is given by all families $\mathscr{F}' = (\mathscr{F}'_{\mathbf{x}'})_{\mathbf{x}' \in \mathbf{X}'}$ of σ-algebras on Ω such that $\mathscr{F}_{\mathbf{x}'_2} = \{\{\omega_2\}, \emptyset\}$ and $\mathscr{F}_{\mathbf{x}'_0} \subseteq \mathscr{F}_{\mathbf{x}'_1}$. The claim follows easily from this. $\square$

Proof of lemma 3.4. Let $\mathscr{F}$, W, and f as in the claim. Further, let $t, u \in \mathbb{T}_f$ with $t < u$ and $E \in \mathscr{F}_{\mathbf{x}_t(f)}$. Then, $\mathbf{x}_t(f) \geq_{\mathbf{X}} \mathbf{x}_u(f)$, whence $E \cap D_{u,f} \in \mathscr{F}_{\mathbf{x}_u(f)}$ by definition 3.1.

The second claim follows from the first because under its hypothesis $\mathscr{F}_{\mathbf{x}_t(f)}$ is a σ-algebra over $D_{t,f} = \Omega$ for all $t \in \mathbb{T}_f$. $\square$

Proof of theorem 3.6. Let $\mathbf{x} \in \mathbf{X}$. Then, by construction, $\mathscr{F}_{\mathbf{x}}$ is a σ-algebra over $D_{\mathbf{x}}$ contained in $\mathscr{A}$, because $\mathscr{G}_{\mathsf{t}(x)}$ is a sub-σ-algebra of $\mathscr{A}$ and all $Y_{\mathbf{x}'}$, ranging over $\mathbf{x}' \in \mathbf{X}$, are $\mathscr{A}$-measurable.

Furthermore, let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{X}$ such that $\mathbf{x}_1 \geq_{\mathbf{X}} \mathbf{x}_2$, and let $E \in \mathscr{F}_{\mathbf{x}_1}$. Hence, there is $E' \in \sigma(Y_{\mathbf{x}'} \mid \mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}_1) \vee \mathscr{G}_{\mathsf{t}(\mathbf{x}_1)}$ such that $E = E' \cap D_{\mathbf{x}_1}$. Note that

$$E' \in \sigma(Y_{\mathbf{x}'} \mid \mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}_1) \vee \mathscr{G}_{\mathsf{t}(\mathbf{x}_1)} \subseteq \sigma(Y_{\mathbf{x}'} \mid \mathbf{x}' \geq_{\mathbf{X}} \mathbf{x}_2) \vee \mathscr{G}_{\mathsf{t}(\mathbf{x}_2)}.$$

As $D_{\mathbf{x}_1} \supseteq D_{\mathbf{x}_2}$, we have $E \cap D_{\mathbf{x}_2} = E' \cap D_{\mathbf{x}_2}$. Hence, $E \cap D_{\mathbf{x}_2} \in \mathscr{F}_{\mathbf{x}_2}$. $\square$

Appendix A.4. Section 4

Proof of lemma 4.1. As $c \cap W_A = \bigcup_{\omega \in A}(c \cap W_\omega)$, we infer from lemma Appendix A.1, part 1, that the following introductory statement holds true:

$$\downarrow (c \cap W_A) = \{x \in F \mid c \cap W_A \supseteq x\} = \{x \in F_A \mid c \supseteq x\} = (\downarrow c) \cap F_A.$$

From this, the claim follows easily: If $x \in P(c \cap W_A)$, then there is $y \in (\downarrow c) \cap F_A$ such that

$$(\dagger) \qquad \uparrow x = \uparrow y \setminus ((\downarrow c) \cap F_A).$$

Hence, $x \in \uparrow y$. There is unique $\omega \in A$ such that $y \in (\downarrow c) \cap T_\omega$. As T_ω is a connected component of $(F, \supseteq)$ and $y \in T_\omega$, we have $\uparrow y \subseteq T_\omega$, whence $x \in T_\omega$ and

$$(*) \qquad \uparrow x = \uparrow y \setminus \downarrow c.$$

If conversely $x \in F_A$ is such that there is $y \in \downarrow c$ satisfying $(*)$, then $x \in \uparrow y$, hence x, y belong to the same connected component, whence $y \in F_A$ and $\uparrow y \subseteq F_A$. We infer that $(\dagger)$ holds true, i.e., $x \in P(c \cap W_A)$ by the introductory statement. $\square$

Lemma Appendix A.3. *Consider the basic version $(F, \pi, \mathbf{X})$ of the simple SDF over W and the exogenous scenario space $(\Omega, \mathscr{A})$ as in subsection 4.3. Let M be the set of maps $\Omega \to \{1, 2\}$. Then the following subsets of W define non-redundant and complete choices:*

- *$c_{f\bullet}$, where $f \in M$;*
- *c_{kg}, where $k = 1, 2$ and $g \in M$;*
- *$c_{\bullet g}$, where $g \in M$.*

The corresponding sets of immediate predecessors are given by $P(c_{f\bullet}) = \operatorname{im} \mathbf{x}_0$; $P(c_{kg}) = \operatorname{im} \mathbf{x}_k$; $P(c_{\bullet g}) = \operatorname{im} \mathbf{x}_1 \cup \operatorname{im} \mathbf{x}_2$.

Proof. Any set c of the form above is non-empty, and as F contains all singletons in $\mathcal{P}(W)$, c is a choice. The sets of immediate predecessors are easily shown to be of the claimed form, using lemma 4.1. From this, we directly infer the non-redundancy and completeness of all the considered choices. $\square$

Proof of lemma 4.3. Let $C = \bigcup_{\mathbf{x}\in\mathbf{X}} \mathscr{C}_{\mathbf{x}}$. By lemma Appendix A.3, all elements of C are non-redundant and complete choices. The statement about the set of immediate predecessors of first $c_{k\bullet}$ and second $c_{\bullet m}$, $k, m \in \{1, 2\}$, stated in lemma Appendix A.3 shows that the first type of choice is available at $\mathbf{x}_0$, while the second is available at $\mathbf{x}_1$ and $\mathbf{x}_2$. □

Proof of lemma 4.4. In view of lemma Appendix A.3, we only have to verify that for any line and corresponding exogenous information structure (EIS) $\mathscr{F}$, and any subset $c \subseteq W$ in that line, all $\mathbf{x} \in \mathbf{X}$ that c is available at and $c_{\mathbf{x}} \in \mathscr{C}_{\mathbf{x}}$, we have

$$\mathbf{x}^{-1}(P(c \cap c_{\mathbf{x}})) \in \mathscr{F}_{\mathbf{x}}.$$

For this, one easily verifies that all $k, k', m' \in \{1, 2\}$ and $f, g \in M$ satisfy:

$$\begin{aligned} c_{f\bullet} \cap c_{k'\bullet} &= c_{k'\bullet} \cap W_{\{f=k'\}}; \\ c_{kg} \cap c_{\bullet m'} &= c_{km'} \cap W_{\{g=m'\}}; \\ c_{\bullet g} \cap c_{\bullet m'} &= c_{\bullet m'} \cap W_{\{g=m'\}}. \end{aligned}$$

As shown in lemma Appendix A.3, we have

$$P(c_{k'\bullet}) = \operatorname{im} \mathbf{x}_0, \quad P(c_{km'}) = \operatorname{im} \mathbf{x}_k, \quad P(c_{\bullet m'}) = \operatorname{im} \mathbf{x}_1 \cup \operatorname{im} \mathbf{x}_2.$$

Then, applying lemma 4.1 and using the definition of $\mathscr{F}$, in each of the five cases respectively, completes the proof. □

Lemma Appendix A.4. *Consider the variant $(F', \pi', \mathbf{X}')$ of the simple SDF over W' and the exogenous scenario space $(\Omega, \mathscr{A})$ as in subsection 4.3. Let M be the set of maps $\Omega \to \{1, 2\}$. Then the following subsets of W' define non-redundant and complete choices:*

- $c'_{f\bullet}$, *where* $f \in M$;
- c'_{kg}, *where* $k = 1, 2$ *and* $g \in M$;
- $c'_{\bullet g}$, *where* $g \in M$.

The corresponding sets of immediate predecessors are given by $P(c'_{f\bullet}) = \operatorname{im} \mathbf{x}'_0$; $P(c'_{kg}) = \operatorname{im} \mathbf{x}'_k$; $P(c'_{\bullet g}) = \operatorname{im} \mathbf{x}'_1 \cup \operatorname{im} \mathbf{x}'_2$.

Proof. Any set c' of the form above is non-empty, and as F' contains all singletons in $\mathcal{P}(W')$, any such c' is a choice. The sets of immediate predecessors are easily shown to be of the claimed form, using lemma 4.1. From this, we directly infer the non-redundancy and completeness of all the considered choices. □

Proof of lemma 4.5. Let $C' = \bigcup_{\mathbf{x}'\in\mathbf{X}'} \mathscr{C}'_{\mathbf{x}'}$. By lemma Appendix A.4, all elements of C' are non-redundant and complete choices. The statement about the set of immediate predecessors of first $c'_{k\bullet}$ and second $c'_{\bullet m}$, $k, m \in \{1, 2\}$, stated in lemma Appendix A.4 shows that the first type of choice is available at $\mathbf{x}'_0$, while the second is available at $\mathbf{x}'_1$ and $\mathbf{x}'_2$. □

Proof of lemma 4.6. In view of lemma Appendix A.4, we only have to verify that for any line and corresponding exogenous information structure (EIS) $\mathscr{F}'$, and any subset $c' \subseteq W$ in that line, all $\mathbf{x}' \in \mathbf{X}'$ that c' is available at and $c'_{\mathbf{x}'} \in \mathscr{C}'_{\mathbf{x}'}$, we have

$$\mathbf{x}'^{-1}(P(c' \cap c'_{\mathbf{x}'})) \in \mathscr{F}'_{\mathbf{x}'}.$$

For this, one easily verifies that all $k, k', m' \in \{1,2\}$ and $f, g \in M$ satisfy:

$$c'_{f\bullet} \cap c'_{k'\bullet} = c'_{k'\bullet} \cap W_{\{f=k'\}};$$
$$c'_{kg} \cap c'_{\bullet m'} = c'_{km'} \cap W_{\{g=m'\}};$$
$$c'_{\bullet g} \cap c'_{\bullet m'} = c'_{\bullet m'} \cap W_{\{g=m'\}}.$$

As shown in lemma Appendix A.4, we have

$$P(c'_{k'\bullet}) = \operatorname{im} \mathbf{x}'_0, \quad P(c'_{km'}) = \operatorname{im} \mathbf{x}'_k, \quad P(c'_{\bullet m'}) = \operatorname{im} \mathbf{x}'_1 \cup \operatorname{im} \mathbf{x}'_2.$$

Then applying lemma 4.1 and using the definition of $\mathscr{F}'$, in each of the three cases respectively, completes the proof. □

Proof of lemma 4.7. Let $t \in \mathbb{T}$ and $c \in \mathscr{C}_t$. Let $x \in F$.

If $x = \{w\}$ for some $w \in W$, then clearly $x \in\downarrow c$ iff $w \in c$. It remains to consider the case $x = x_u(w)$ for some $w = (\omega, f) \in W$ and $u \in \mathbb{T}$.

If $t < u$ and $w \in c$, then any $w' = (\omega', f') \in x$ satisfies $\omega' = \omega$ and $f'|_{[0,u)_\mathbb{T}} = f|_{[0,u)_\mathbb{T}}$, in particular $f'|_{[0,t]} = f|_{[0,t]}$, hence $w' \in c$. We conclude for this case that $x \subseteq c$.

If $t < u$ and $w \in c$ do not both hold true, then either $w \notin c$, or $w \in c$ and $u \leq t$. If $w \notin c$, then $w \in x \setminus c$, hence $x \nsubseteq c$. If $w \in c$ and $u \leq t$, then, by assumption AP.C1, there is $w' \in x_t(w) \setminus c \subseteq x_u(w) \setminus c = x \setminus c$. Whence $x \nsubseteq c$. □

Lemma Appendix A.5. *Consider an action path* SDF $(F, \pi, \mathbf{X})$ *over an exogenous scenario space* $(\Omega, \mathscr{A})$*, and let* $W = \bigcup F$*. Let* $t \in \mathbb{T}$ *and* $w \in W$*. Then,*

$$\uparrow x_t(w) = \{x_u(w) \mid u \in \mathbb{T}\colon u \leq t\}.$$

Proof. The inclusion $\supseteq$ is clear from the definition of the nodes of the action path SDF. For the converse inclusion, let $x \in\uparrow x_t(w)$. If x is a singleton, then $x = x_t(w)$. If x is no singleton, then it is a move and there are $u \in \mathbb{T}$ and $w' \in W$ with $x = x_u(w')$. As $w \in x_t(w) \subseteq x$, we even have $x = x_u(w)$. By means of lemma 2.13, we infer $u = \mathfrak{t}(x) \leq \mathfrak{t}(x_t(w)) = t$. □

Proof of lemma 4.8. Let $t \in \mathbb{T}$ and $c \in \mathscr{C}_t$. Let $x \in F$. By definition, $x \in P(c)$ is equivalent to the existence of $y \in\downarrow c$ satisfying

$$(*) \qquad \uparrow x = \uparrow y \setminus \downarrow c.$$

If $x = x_t(w)$ for $w \in c$, then, by assumption AP.C1, there is $w' \in x_t(w) \setminus c$. Thus, x has at least two elements. In view of lemmata 4.7 and Appendix A.5 and assumption AP.W1, $(*)$ is satisfied for $y = \{w\}$.

If, conversely, $(*)$ is satisfied for some $y \in F$, then x is a move, whence the existence of $w_0 \in W$ and $t_0 \in \mathbb{T}$ with $x = x_{t_0}(w_0)$. Moreover, there is $w \in y$, and as $y \in\downarrow c$, we obtain $w \in c$. By $(*)$, we know that $x \in\uparrow y$, hence also $w \in x$. Thus, $x = x_{t_0}(w)$. From representation $(*)$ and lemma 4.7, we immediately get $t_0 \leq t$. Moreover, by assumption AP.C1 on $c \in \mathscr{C}_t$, $x_t(w)$ has at least two elements. Indeed, as $w \in c$, there is $w' \in x_t(w) \setminus c$. Hence, by assumption AP.W1, $x_t(w) \neq x_u(w)$ for all $u > t$. Hence, $x_t(w) \in\uparrow y \setminus \downarrow c = \uparrow x$, again by representation $(*)$ and lemma 4.7. Thus, by lemma 2.13, $t = \mathfrak{t}(x_t(w)) \leq \mathfrak{t}(x) = t_0$. Hence, $t_0 = t$ and $x = x_t(w)$ with $w \in c$. □

Proof of lemma 4.9. By assumption AP.C0, c is a non-empty union of singletons in $\mathcal{P}(W)$, which are elements of F by construction. Hence, it is a choice.

Concerning non-redundancy, let $\omega \in \Omega$ and suppose there is $w \in c \cap W_\omega$. Then, $x_t(w) \in P(c) \cap T_\omega$, by lemma 4.8. Hence, by contraposition, if $P(c) \cap T_\omega = \emptyset$, then $c \cap W_\omega = \emptyset$ as well.

Concerning completeness, let $\mathsf{x} \in \mathbf{X}$ such that there is $\omega \in D_{\mathsf{x}}$ with $\mathsf{x}(\omega) \in P(c)$. By lemma 4.8, there is $f \in \mathbb{A}^{\mathbb{T}}$ such that $(\omega, f) \in c$ and $\mathsf{x}(\omega) = x_t(\omega, f)$. First, we infer that $D_{t,f} \neq \emptyset$, whence $D_{\mathsf{x}} = D_{t,f}$ and $\mathsf{x} = \mathsf{x}_t(f)$, by proposition 2.3. Second, we infer that $x_t(\omega, f) \cap c \neq \emptyset$. By assumption AP.C2, for any $\omega' \in D_{t,f}$ there is $w' \in x_t(\omega', f) \cap c$. In particular, $\mathsf{x}(\omega') = x_t(\omega, f) = x_t(w')$. Hence, $\mathsf{x}(\omega') \in P(c)$, by lemma 4.8. We conclude that $\mathsf{x}^{-1}(P(c)) = D_{\mathsf{x}}$.

Regarding the proof of the second sentence, let $\mathsf{x} \in \mathbf{X}$ such that c is available at x. There is $\omega \in D_{\mathsf{x}}$. Then, $\mathsf{x}(\omega) \in P(c)$. As in the proof of completeness above, we infer the existence of $f \in \mathbb{A}^{\mathbb{T}}$ such that $(\omega, f) \in c$, $\omega \in D_{t,f} = D_{\mathsf{x}}$ and $\mathsf{x} = \mathsf{x}_t(f)$. ☐

Proof of proposition 4.10. Let $\mathsf{x} \in \mathbf{X}$, $t = \mathsf{t}(\mathsf{x})$, and $c \in \mathscr{C}_{\mathsf{x}}$. In particular, $c \in \mathscr{C}_t$, and, by lemma 4.9, c is a complete and non-redundant choice. Let $\omega \in D_{\mathsf{x}}$. Then, by definition of $\mathscr{C}_{\mathsf{x}}$, there is $w \in \mathsf{x}(\omega) \cap c$. Hence, $\mathsf{x}(\omega) = x_{\mathsf{t}(x)}(w) = x_t(w)$, and, by lemma 4.8, $\mathsf{x}(\omega) \in P(c)$. We conclude that c is available at x. ☐

Proof of theorem 4.11. (First two statements): Let $c = c(A_{<t}, i, g)$. By lemma 4.9, c is a non-redundant and complete choice which proves the first claim.

Next, let $\mathsf{x} \in \mathbf{X}$ such that c is available at x. By definition of $\mathbf{X}$, there is $f_0 \in \mathbb{A}^{\mathbb{T}}$ with $D_{t,f_0} \neq \emptyset$ such that $\mathsf{x} = \mathsf{x}_t(f_0)$ which we fix for the entire proof.

The second claim of the theorem is also proven directly. To show this, let $\omega \in D_{\mathsf{x}}$. Then $\mathsf{x}(\omega) \in P(c)$, hence, by lemma 4.8, there is $w \in c$ with $\mathsf{x}(\omega) = x_t(w)$. There is $f \in \mathbb{A}^{\mathbb{T}}$ such that $w = (\omega, f)$, and as $w \in c$, we have $f(t) \in A^{i,g}_{t,\omega}$. Hence, $\omega \in D$. We conclude that $D_{\mathsf{x}} \subseteq D$.

(Helpful statements for $c' \in \mathscr{C}^i_{\mathsf{x}}$): Let $c' \in \mathscr{C}^i_{\mathsf{x}}$. We compute the set $P(c \cap c')$ and its preimage under x. By definition of $\mathscr{C}^i_{\mathsf{x}}$, there are $A'_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ and $A'^i_t \in \mathscr{B}(\mathbb{A}^i)$ such that, with $A'_t = (A'_{t,\omega})_{\omega \in \Omega}$ and $A'_{t,\omega} = (p^i)^{-1}(A'^i_t)$ for all $\omega \in D_{\mathsf{x}}$, and $A'_{t,\omega} = \emptyset$ for all $\omega \notin D_{\mathsf{x}}$, we have $c' = c(A'_{<t}, A'_t)$, $c' \in \mathscr{C}_t$, and, for all $\omega \in D_{\mathsf{x}}$, $\mathsf{x}(\omega) \cap c' \neq \emptyset$.

Let $c_0 = c(A_{<t} \cap A'_{<t}, i, g)$. By definition of c and c', we have

$$\begin{aligned}
&c \cap c' \\
&= \{(\omega, f) \in W_{D_{\mathsf{x}}} \mid f|_{[0,t)_{\mathbb{T}}} \in A_{<t} \cap A'_{<t},\ p^i \circ f(t) = g(\omega) \in A'^i_t\} \\
&= c(A_{<t} \cap A'_{<t}, i, g) \cap \{(\omega, f) \in W_{D_{\mathsf{x}}} \mid g(\omega) \in A'^i_t\} \\
&= c_0 \cap W_{g|^{-1}_{D_{\mathsf{x}}}(A'^i_t)}.
\end{aligned}$$

Using lemma 4.1, we infer

$$(*) \qquad P(c \cap c') = P(c_0) \cap F_{g|^{-1}_{D_{\mathsf{x}}}(A'^i_t)}.$$

Next, we show that

$$(\circ) \qquad \forall \omega \in D_{\mathsf{x}} \exists f \in \mathbb{A}^{\mathbb{T}}: \quad (\omega, f) \in c_0,\ \mathsf{x}(\omega) = x_t(\omega, f).$$

For the proof of $(\circ)$, let $\omega \in D_{\mathsf{x}}$. As x is available both at c and c', we have $\mathsf{x}(\omega) = x_t(\omega, f_0) \in P(c) \cap P(c')$. Hence, by lemma 4.8 applied to $c, c' \in \mathscr{C}_t$, there are $f, f' \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, f) \in c$ and $(\omega, f') \in c'$ such that

$$f'|_{[0,t)_{\mathbb{T}}} = f_0|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}}.$$

Thus, $f|_{[0,t)_{\mathbb{T}}} \in A_{<t} \cap A'_{<t}$. By definition of f, we must have $p^i \circ f(t) = g(\omega)$. Hence, $(\omega, f) \in c_0$. By definition of f, we also have $\mathsf{x}(\omega) = x_t(\omega, f)$.

We infer that $c_0 \in \mathscr{C}_t$. Indeed, assumption AP.C0 is satisfied by $(\circ)$. Concerning AP.C1, note that $c_0 \subseteq c$. Let $w \in c_0$. Thus $w \in c$. Hence, by AP.C1 applied to c, there is $w' \in x_t(w) \setminus c \subseteq$

$x_t(w) \setminus c_0$. Regarding AP.C2, let $f \in \mathbb{A}^{\mathbb{T}}$ with $f|_{[0,t)_{\mathbb{T}}} \in A_{<t} \cap A'_{<t}$ such that there is $\omega \in D_{t,f}$ satisfying

$$x_t(\omega, f) \cap c_0 \neq \emptyset.$$

Let $\omega' \in D_{t,f}$. As $c_0 \subseteq c$, we infer that $x_t(\omega, f) \cap c \neq \emptyset$. Hence, by AP.C2 applied to c, there exists $w' \in x_t(\omega', f) \cap c$. Let $f' \in \mathbb{A}^{\mathbb{T}}$ such that $w' = (\omega', f')$. Then, by definition of w' and f', we get $f'|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}} \in A_{<t} \cap A'_{<t}$, $\omega' \in D$, and $p^i(f'(t)) = g(\omega')$. Hence, $(\omega', f') \in x_t(\omega', f) \cap c_0$. We conclude that $c_0 \in \mathscr{C}_t$.

Hence, by lemma 4.9, c_0 defines a non-redundant and complete choice. Moreover, by ($\circ$) and lemma 4.8, c_0 is available at $\mathbf{x}$. In particular, we have $\mathbf{x}^{-1}(P(c_0)) = D_{\mathbf{x}}$. By ($*$), we get:

$$(\dagger) \qquad \mathbf{x}^{-1}(P(c \cap c')) = \mathbf{x}^{-1}(P(c_0)) \cap \mathbf{x}^{-1}(F_{g|_{D_{\mathbf{x}}}^{-1}(A_t'^i)}) = D_{\mathbf{x}} \cap g|_{D_{\mathbf{x}}}^{-1}(A_t'^i) = g|_{D_{\mathbf{x}}}^{-1}(A_t'^i).$$

(First implication): If $g|_{D_{\mathbf{x}}}$ is $\mathscr{F}^i_{\mathbf{x}}$-measurable, then, by ($\dagger$), $\mathbf{x}^{-1}(P(c \cap c')) \in \mathscr{F}^i_{\mathbf{x}}$ for all $c' \in \mathscr{C}^i_{\mathbf{x}}$.

(Second implication): Suppose conversely that $\mathbf{x}^{-1}(P(c \cap c')) \in \mathscr{F}^i_{\mathbf{x}}$ for all $c' \in \mathscr{C}^i_{\mathbf{x}}$, and that assumption AP.C3 is satisfied. Hence, there is $A'_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ with $f|_{[0,t)_{\mathbb{T}}} \in A'_{<t}$ for all $(\omega, f) \in c$, and an intersection-stable generator $\mathscr{G}(\mathbb{A}^i)$ of $\mathscr{B}(\mathbb{A}^i)$ such that for all $G \in \mathscr{G}(\mathbb{A}^i)$, we have $c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}^i_{\mathbf{x}} \cup \{\emptyset\}$. Let $G \in \mathscr{G}(\mathbb{A}^i)$ and $c' = c(A'_{<t}, A_t^{i,G})$.

Assume first that $c' = \emptyset$. Let $\omega \in D_{\mathbf{x}}$. Then, as c is available at $\mathbf{x}$, $\mathbf{x}(\omega) \in P(c)$. Hence, by lemma 4.8, there is $f \in \mathbb{A}^{\mathbb{T}}$ with $w = (\omega, f) \in \mathbf{x}(\omega) \cap c$. In particular, using the definition of c and property AP.C3(1) of $A'_{<t}$, $p^i \circ f(t) = g(\omega)$ and $f|_{[0,t)_{\mathbb{T}}} \in A'_{<t}$. As $c' = \emptyset$, $g(\omega) \notin G$. Hence, $g|_{D_{\mathbf{x}}}^{-1}(G) = \emptyset \in \mathscr{F}^i_{\mathbf{x}}$.

Second, assume that $c' \neq \emptyset$. Then, by definition, $c' \in \mathscr{C}^i_{\mathbf{x}}$. Represent c' as in the beginning of the "helpful statements" part above, with $A_t'^i = G$. Let c_0 be defined as in the "helpful statements" part.

Then, we can use these helpful statements to infer that c_0 is a non-redundant and complete choice available at $\mathbf{x}$. Hence, $\mathbf{x}^{-1}(P(c_0)) = D_{\mathbf{x}}$, and thus ($\dagger$) and the hypothesis imply that

$$(g|_{D_{\mathbf{x}}})^{-1}(G) = \mathbf{x}^{-1}(P(c \cap c')) \in \mathscr{F}^i_{\mathbf{x}}.$$

We conclude that $(g|_{D_{\mathbf{x}}})^{-1}(G) \in \mathscr{F}^i_{\mathbf{x}}$ for all $G \in \mathscr{G}(\mathbb{A}^i)$. As $\mathscr{G}(\mathbb{A}^i)$ is an intersection-stable generator of $\mathscr{B}(\mathbb{A}^i)$, $g|_{D_{\mathbf{x}}}$ is $\mathscr{F}^i_{\mathbf{x}}$-measurable. $\square$

Proofs for example 4.12. (Ad $W = \Omega \times \mathbb{A}^{\mathbb{T}}$): We show the following more general statement: If $t \in \mathbb{T}$ and $A_{<t} \subseteq \mathbb{A}^{[0,t)_{\mathbb{T}}}$ is non-empty, and moreover $A_t = (A_{t,\omega})_{\omega \in \Omega} \in \mathcal{P}(\mathbb{A})^{\Omega}$ is such that

$$(*) \qquad \forall \omega \in \Omega\colon \quad \emptyset \subsetneq A_{t,\omega} \subsetneq \mathbb{A},$$

then $c(A_{<t}, A_t) \in \mathscr{C}_t$. This includes the two following cases: A) $A_t = A_t^{i,g}$ for given $i \in I$ and a map $g\colon \Omega \to \mathbb{A}^i$ because p^i is surjective and $\mathbb{A}^i$ has at least two elements; and B) $A_{t,\omega} = (p^i)^{-1}(G)$ for some $i \in I$ and some set $\emptyset \subsetneq G \subsetneq \mathbb{A}^i$, again because p^i is surjective.

As $A_{<t} \neq \emptyset$ and ($*$) is assumed, there is $w = (\omega, f) \in W = \Omega \times \mathbb{A}^{\mathbb{T}}$ such that $f|_{[0,t)_{\mathbb{T}}} \in A_{<t}$ and $f(t) \in A_{t,\omega}$. Hence, $w \in c(A_{<t}, A_t)$. Thus, AP.C0 is satisfied.

Regarding AP.C1, let $w = (\omega, f) \in c(A_{<t}, A_t)$. There is $f' \in \mathbb{A}^{\mathbb{T}}$ with $f'|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}}$ and $f'(t) \notin A_{t,\omega}$ (by ($*$)). Then $w' = (\omega, f') \in \Omega \times \mathbb{A}^{\mathbb{T}} = W$, hence $w' \in x_t(w)$, but $w' \notin c(A_{<t}, A_t)$.

Regarding AP.C2, let $f \in \mathbb{A}^{\mathbb{T}}$ with $f|_{[0,t)_{\mathbb{T}}} \in A_{<t}$ and $\omega \in D_{t,f}$. There is $f' \in \mathbb{A}^{\mathbb{T}}$ with $f'|_{[0,t)_{\mathbb{T}}} = f|_{[0,t)_{\mathbb{T}}}$ and $f'(t) \in A_{t,\omega}$, by ($*$). As $(\omega, f') \in \Omega \times \mathbb{A}^{\mathbb{T}} = W$, we infer

$$(\omega, f') \in x_t(\omega, f) \cap c(A_{<t}, A_t).$$

Hence, $c(A_{<t}, A_t) \in \mathscr{C}_t$. The general statement above is proven. By considering the case A), we infer that for all $i \in I$ and $g\colon \Omega \to \mathbb{A}^i$, $c(A_{<t}, i, g) \in \mathscr{C}_t$.

Regarding AP.C3, let $\mathbf{x} \in \mathbf{X}$ be such that, with $t = \mathsf{t}(\mathbf{x})$, there are $D \in \mathscr{A}$, $g\colon D \to \mathbb{A}^i$ and $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ satisfying $c(A_{<t}, i, g) \in \mathscr{C}_t$ and making $c(A_{<t}, i, g)$ available at $\mathbf{x}$. By lemma 4.9, there is $(\omega_0, f_0) \in c(A_{<t}, i, g)$ such that $\mathbf{x} = \mathbf{x}_t(f_0)$ and $\omega_0 \in D_\mathbf{x}$.

Let $A'_{<t} = \mathbb{A}^{[0,t)_\mathbb{T}}$ and $\mathscr{G}(\mathbb{A}^i) = \mathscr{B}(\mathbb{A}^i) \setminus \{\mathbb{A}^i\}$ which is obviously an intersection-stable generator of $\mathscr{B}(\mathbb{A}^i)$. Clearly, $A_{<t} \subseteq A'_{<t}$ so that property AP.C3(1) is satisfied. Let $G \in \mathscr{G}(\mathbb{A}^i)$. If $G = \emptyset$, then $c(A'_{<t}, A_t^{i,G}) = \emptyset$. If $G \in \mathscr{G}(\mathbb{A}^i) \setminus \{\emptyset\}$, then the general statement above in case B) applies, hence $c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}_t$. To complete the proof that $c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}_\mathbf{x}$, it remains to prove property (4) in the definition of $\mathscr{C}_\mathbf{x}$, (3) having been proven just before and (1) and (2) being evident. For this, let $\omega' \in D_\mathbf{x}$. Then there is $f' \in \mathbb{A}^\mathbb{T}$ such that $f'|_{[0,t)_\mathbb{T}} = f_0|_{[0,t)_\mathbb{T}} \in \mathbb{A}^{[0,t)_\mathbb{T}} = A'_{<t}$ and $p^i \circ f'(t) \in G$, because G is non-empty and p^i surjective. Then, $w' = (\omega', f') \in \Omega \times \mathbb{A}^\mathbb{T} = W$, and thus $w' \in \mathbf{x}_t(f_0)(\omega') \cap c(A'_{<t}, A_t^{i,G})$. Property (4) therefore holds true and the proof of the statement $c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}_\mathbf{x} \cup \{\emptyset\}$ is complete.

(Ad timing game): Let $t \in \mathbb{T}$. Further, let $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ be a non-empty set of component-wise decreasing paths f_t such that $p^i \circ f_t = 1_{[0,t)_\mathbb{T}}$, and $g\colon \Omega \to \{0,1\} = \mathbb{A}^i$. Let $c = c(A_{<t}, i, g)$. We are going to prove that $c \in \mathscr{C}_t$.

Regarding AP.C0, let $\omega \in \Omega$. By assumption on $A_{<t}$, there is component-wise decreasing $f\colon \mathbb{T} \to \mathbb{A}$ such that $f|_{[0,t)_\mathbb{T}} \in A_{<t}$ and $p^i \circ f(t) = g(\omega)$. Hence, $(\omega, f) \in W$, and even $(\omega, f) \in c$. Thus, $c \neq \emptyset$.

Regarding AP.C1, let $w = (\omega, f) \in c$. There is component-wise decreasing $f' \in \mathbb{A}^\mathbb{T}$ with $f'|_{[0,t)_\mathbb{T}} = f|_{[0,t)_\mathbb{T}}$ and $p^i \circ f'(t) \neq g(\omega)$, because $p^i \circ f|_{[0,t)_\mathbb{T}} = 1_{[0,t)_\mathbb{T}}$. Then $w' = (\omega, f') \in W$, hence $w' \in x_t(w)$, but $w' \notin c$.

Regarding AP.C2, let $f \in \mathbb{A}^\mathbb{T}$ with $f|_{[0,t)_\mathbb{T}} \in A_{<t}$ and $\omega \in D_{t,f}$. There is component-wise decreasing $f' \in \mathbb{A}^\mathbb{T}$ with $f'|_{[0,t)_\mathbb{T}} = f|_{[0,t)_\mathbb{T}}$ and $p^i \circ f'(t) = g(\omega)$, by assumption on $A_{<t}$. As $(\omega, f') \in W$, we infer

$$(\omega, f') \in x_t(\omega, f) \cap c.$$

We conclude that $c \in \mathscr{C}_t$.

Regarding assumption AP.C3, let $\mathbf{x} \in \mathbf{X}$ be such that, with $t = \mathsf{t}(\mathbf{x})$, there are $D \in \mathscr{A}$, $g\colon D \to \mathbb{A}^i = \{0,1\}$ and $A_{<t} \subseteq \mathbb{A}^{[0,t)_\mathbb{T}}$ satisfying $c(A_{<t}, i, g) \in \mathscr{C}_t$ and making $c(A_{<t}, i, g)$ available at $\mathbf{x}$. By lemma 4.9, there is $(\omega_0, f_0) \in c(A_{<t}, i, g)$ such that $\omega_0 \in D_{t,f_0} = D_\mathbf{x}$ and $\mathbf{x} = \mathbf{x}_t(f_0)$. As $c(A_{<t}, i, g) \in \mathscr{C}_t$, assumption AP.C1 implies, that $p^i \circ f_0|_{[0,t)_\mathbb{T}}$ is constant with value 1. Indeed, there is $f_1 \in \mathbb{A}^\mathbb{T}$ with $(\omega, f_1) \in x_t(\omega, f_0) \setminus c(A_{<t}, i, g)$. In particular, f_1 is component-wise decreasing and takes the same values on $[0,t)_\mathbb{T}$ as f_0. If $p^i \circ f_0$ took the value 0 at some point in $[0,t)_\mathbb{T}$, the monotonicity would imply $p^i \circ f_1(t) = 0 = p^i \circ f_0(t) = g(\omega)$, whence $(\omega, f_1) \in c(A_{<t}, i, g)$, in contradiction to the choice of f_1.

Let $A'_{<t}$ be the set of component-wise decreasing $f_t \in \mathbb{A}^{[0,t)_\mathbb{T}}$ such that $p^i \circ f_t$ is constant with value 1. As $A_{<t} \subseteq A'_{<t}$, property AP.C3(1) is clearly satisfied. Further, let $\mathscr{G}(\mathbb{A}^i) = \mathscr{B}(\mathbb{A}^i) \setminus \{\mathbb{A}^i\}$ which is obviously an intersection-stable generator of $\mathscr{B}(\mathbb{A}^i)$. As $\mathbb{A}^i = \{0,1\}$, $\mathscr{G}(\mathbb{A}^i) = \{\{0\}, \{1\}, \emptyset\}$. Let $G \in \mathscr{G}(\mathbb{A}^i)$. If $G = \emptyset$, then $c(A'_{<t}, A_t^{i,G}) = \emptyset$. Else, that is, if G is a singleton, then $c(A'_{<t}, A_t^{i,G}) = c(A'_{<t}, i, g_G)$ for the constant map g_G with value given by the unique element of G. We have shown just beforehand that $c(A'_{<t}, i, g_G) \in \mathscr{C}_t$. To complete the proof of the fact that $c(A'_{<t}, A_t^{i,G}) \in \mathscr{C}_\mathbf{x}$, it thus remains to show axiom (4) in the definition of $\mathscr{C}_\mathbf{x}$, because (3) has just been proven and (1) and (2) are evident by construction. For this, let $\omega' \in D_\mathbf{x}$. As $p^i \circ f_0$ only takes the value 1 on $[0,t)_\mathbb{T}$, there is component-wise decreasing $f'\colon \mathbb{T} \to \mathbb{A}$ such that $f'|_{[0,t)_\mathbb{T}} = f_0|_{[0,t)_\mathbb{T}}$ and $p^i \circ f'(t) = g_G(\omega')$.

Hence, $(\omega', f') \in \mathbf{x}_t(f_0)(\omega') \cap c(\tilde{A}'_{<t}, i, g_G) = \mathbf{x}(\omega') \cap c(A'_{<t}, A_t^{i,G})$. This completes the proofs for the timing game example.

(Ad up-and-out option exercise example): Let $t \in \mathbb{R}_+$, $A_{<t} = \{1\}^{\mathbb{R}_+}$ and D be the set of $\omega \in \Omega$ such that $\max_{u\in[0,t]} P_u(\omega) < 2$. We suppose that $D \neq \emptyset$. Let $g\colon D \to \{0,1\}$ be a map and let $c = c(A_{<t}, i, g)$. We are going to prove that $c \in \mathscr{C}_t$.

We start with the proof of AP.C0. There is $\omega_0 \in D$. As P is continuous, there is $\varepsilon > 0$ such that $\max_{u\in[0,t+\varepsilon]} P_u(\omega_0) < 2$. Then, regardless of the value of $g(\omega_0)$, there is decreasing $f\colon \mathbb{R}_+ \to \mathbb{A}$ such that $f|_{[0,t)} = 1_{[0,t)}$, $f(t) = g(\omega_0)$, and $f(t+\varepsilon) = 0$. Hence, $(\omega_0, f) \in W$, and even $(\omega_0, f) \in c$. Thus, $c \neq \emptyset$.

Regarding AP.C1, let $w = (\omega, f) \in c$. Then, $\max_{u\in[0,t]} P_u(\omega) < 2$, and by continuity of P, there is $\varepsilon > 0$ such that $\max_{u\in[0,t+\varepsilon]} P_u(\omega) < 2$. Hence, regardless of the value of $g(\omega)$, there is decreasing $f' \in \mathbb{A}^{\mathbb{R}_+}$ with $f'|_{[0,t)} = f|_{[0,t)} = 1_{[0,t)}$, $f'(t) \neq g(\omega)$, and $f'(t+\varepsilon) = 0$. Then $w' = (\omega, f') \in W$, hence $w' \in x_t(w)$, but $w' \notin c$.

Regarding AP.C2, let $f \in \mathbb{A}^{\mathbb{R}_+}$ with $f|_{[0,t)} \in A_{<t}$ and $\omega \in D_{t,f}$. Then there is decreasing $\tilde{f}\colon \mathbb{R}_+ \to \mathbb{A}$ with $(\omega, \tilde{f}) \in x_t(\omega, f)$, that is, $\tilde{f}|_{[0,t)} = f|_{[0,t)}$ and $\max_{u\in[0,t]} P_u(\omega) < 2$. By continuity of P, there is $\varepsilon > 0$ such that $\max_{u\in[0,t+\varepsilon]} P_u(\omega) < 2$. Hence, regardless of the value of $g(\omega)$, there is decreasing $f' \in \mathbb{A}^{\mathbb{R}_+}$ with $f'|_{[0,t)} = f|_{[0,t)} = 1_{[0,t)}$, $f'(t) = g(\omega)$, and $f'(t+\varepsilon) = 0$. Hence, $(\omega, f') \in W$, and we infer

$$(\omega, f') \in x_t(\omega, f) \cap c.$$

We conclude that $c \in \mathscr{C}_t$.

Regarding assumption AP.C3, let $\mathbf{x} \in \mathbf{X}$ be such that, with $t = \mathsf{t}(\mathbf{x})$, there are $\tilde{D} \in \mathscr{A}$, $g\colon \tilde{D} \to \mathbb{A}^i = \{0,1\}$ and $A_{<t} \subseteq \mathbb{A}^{[0,t)}$ satisfying $c(A_{<t}, i, g) \in \mathscr{C}_t$ and making $c(A_{<t}, i, g)$ available at $\mathbf{x}$. By lemma 4.9, there is $(\omega_0, f_0) \in c(A_{<t}, i, g)$ such that $\mathbf{x} = \mathbf{x}_t(f_0)$ and $\omega_0 \in D_{\mathbf{x}}$. As $c(A_{<t}, i, g) \in \mathscr{C}_t$, assumption AP.C1 implies, that $f_0|_{[0,t)}$ is constant with value 1. Indeed, there is $f_1 \in \mathbb{A}^{\mathbb{T}}$ with $(\omega, f_1) \in x_t(\omega, f_0) \setminus c(A_{<t}, i, g)$. In particular, f_1 is component-wise decreasing and takes the same values on $[0,t)_{\mathbb{T}}$ as f_0. If f_0 took the value 0 at some point in $[0,t)$, the monotonicity would imply $f_1(t) = 0 = f_0(t) = g(\omega)$, whence $(\omega, f_1) \in c(A_{<t}, i, g)$, in contradiction to the choice of f_1.

Let $A'_{<t} = \{1\}^{[0,t)}$. As $A_{<t} \subseteq A'_{<t}$, property AP.C3(1) is clearly satisfied. Further, let $\mathscr{G}(\{0,1\}) = \mathscr{B}(\{0,1\}) \setminus \{\{0,1\}\}$ which is obviously an intersection-stable generator of $\mathscr{B}(\{0,1\})$. Note that $\mathscr{G}(\{0,1\}) = \{\{0\},\{1\},\emptyset\}$. Let $G \in \mathscr{G}(\{0,1\})$. If $G = \emptyset$, then $c(A'_{<t}, A_t^{i,G}) = \emptyset$. Else, that is, if G is a singleton, then $c(A'_{<t}, A_t^{i,G}) = c(A'_{<t}, i, g_G)$ for the constant map g_G with value given by the unique element of G.

We have shown just above that $c(A'_{<t}, i, g_G) \in \mathscr{C}_t$ because $A'_{<t} \subseteq \mathbb{A}^{[0,t)}$ is a non-empty set of decreasing paths f_t such that $f_t = 1_{[0,t)}$. Hence, $c(A'_{<t}, i, g_G) \in \mathscr{C}_t$. To complete the proof of the fact that $c(A'_{<t}, i, g_G) \in \mathscr{C}_{\mathbf{x}}$, it remains to show axiom (4) in the definition of $\mathscr{C}_{\mathbf{x}}$, because (3) has just been proven and (1) and (2) are evident by construction. For this, let $\omega' \in D_{\mathbf{x}} = D_{t,f_0}$. In particular, $\max_{u\in[0,t]} P_u(\omega') < 2$. By continuity of P, there is $\varepsilon > 0$ such that $\max_{u\in[0,t+\varepsilon]} P_u(\omega') < 2$. Then, regardless of the value of $g(\omega)$, and as f_0 only takes the value 1 on $[0,t)$, there is component-wise decreasing $f'\colon \mathbb{R}_+ \to \mathbb{A}$ such that $f'|_{[0,t)} = f_0|_{[0,t)}$, $f'(t) = g_G(\omega')$, and $f'(t+\varepsilon) = 0$. Hence, $(\omega', f') \in \mathbf{x}(\omega') \cap c(\tilde{A}'_{<t}, i, g_G)$. This completes the proofs for the up-and-out option exercise problem example. $\square$