

The choice environment, constraints, and rational procedures

Paulo Oliva

School of Electronic Engineering and Computer Science, Queen Mary University London

Philipp Zahn

Department of Economics, University of St. Gallen

September 19, 2018

Abstract

In decision theory a choice problem is abstractly represented as a set of alternatives. But when facing a concrete choice, an agent typically cannot access this set at once. For instance, when buying online, he needs to navigate a search result in multiple steps, manually clicking to move from page to page. An agent's choice procedure has to respect this constraint.

In this paper we develop a theory of constraints imposed by the environment and of choice procedures operating on them. We ask which properties procedures have to fulfill such that their choices can equivalently be understood as a maximization of a rational preference relation when the constraints are ignored. We fully characterize such procedures and establish that they are linked to well known properties such as Sen's α . What is more, we show that there exist purely procedural properties that are sufficient for a rationalization.

JEL Classification: D00 D01

1 Introduction

Suppose you want to buy an external hard drive and you go to an online retailer. A search returns a page containing a list of 10 alternatives, plus the option to explore the next pages containing more alternatives. Before you can explore item 11, you need to navigate to the next page. Item 21 will be two clicks away from your initial search. The way the alternatives are presented to you are constraints on your decision procedure. And constraints like this are ubiquitous. Similar to budgets and prices, which restrict *what* you can choose, constraints in the choice environment restrict *how* you can choose. This paper focuses on the modeling of the choice environment, and the corresponding restrictions they impose on decision procedures.

Our suggestion is to think of different environments as *representation languages*, and as such, to model these using *context-free grammars* (CFG). Any such grammar determines how valid sentences can be formed on the basis of some primitive "constructors". It thereby also determines how a sentence is deconstructed or parsed. Applied to our setting, the constructors determine how problem representations are built from a set of alternatives. From the agent's perspective, the grammar constructors restrict how a choice procedure can navigate the representation and access the alternatives.

Consider again the hard drive example. Let X be the set of all hard drives. The context-free grammar can be represented as (informally for now):

Item $\Rightarrow x$, for each $x \in X$

Page $\Rightarrow [\text{Item}_1, \dots, \text{Item}_{10}]$

Result $\Rightarrow \text{Page} \mid \text{Page}, \text{Result}$

Each line is a rule for generating words in the language, and the vertical bar in the "Result" line indicates a choice: a Result is either a single Page, or it is a Page followed by more Results.

This simple example already leads to several interesting observations. First, a problem representation is often a nesting of different sub-representations. We have the single items, then lists of 10 such items forming a page, and finally a sequence of pages forming a search result. Secondly, a search result is defined recursively: Results can be formed of a Page together with more Results. This naturally models scenarios where the set of alternatives is so large that it can be thought of as infinite (think of a Google search). Thirdly, when presented with a result, an agent can either go through the first 10 results on the initial page, *or* he can access the next part of the search results. The grammar represents these two access points by the two constructors for Page and Result. A choice procedure is a sequence of operations on these constructors. For instance, an agent who always chooses the first alternative of the search results picks the first page of the result, and the first item on that page. In a *satisficing* procedure, in contrast, an agent picks the first page and then goes through each of the ten items on that page until a satisfactory item is found. If none is found on the first page, the agent navigates to the remaining search result and repeats this procedure.

We investigate how procedures operating on represented problems are related to classical choice theory. Specifically, we consider (i) the case where the choices implied by a procedure form a (single-valued) choice function and (ii) the case where a decision procedure gives rise to a multi-valued choice correspondence. We characterize when the implied choice function, respectively choice correspondence, is rationalized by a preference relation, by adapting Sen's α and γ conditions [13] to our framework.

To assess whether an agent chooses rationally we have to compare his choices over different choice problems. This is in spirit of classical choice theory. We also ask whether there are properties of the choice process itself which guarantee rationalizability. To this end, we introduce two conditions which, if they jointly hold, are sufficient. The first condition is a *divide and conquer* strategy: The agent reduces a larger decision problem to smaller decision problems and first decides on them before combining the sub-decisions. This principle is prominent in computer science but has been introduced into economic choice theory by [8] as *path independence*. Importantly, whether an agent can proceed in this way, depends on the structure of the environment. The second condition is that, at each step of this reduction via divide and conquer, the agent fulfills a (weak) version of an independence of irrelevant alternatives condition. If these two conditions hold, then choice can be rationalized. Thus, knowing the *process* by which the agent chooses can be sufficient to tell whether his choices can be rationalized.

What insights do we get apart from the results above? Being used to budget constraints, it is tempting to view the restrictions from the environment as a burden only which the agent has to respect. However, the constructors may carry relevant economic information. For example, when buying a hard drive, alternatives are normally sorted according to price, capacity or average rating. An agent, aware of this, may choose the first element of the list if he only cares about one of these criteria. Thus, the environment can also serve as a resource that allows the agent to simplify his procedure.

The resourcefulness of the environment is influenced by the nesting of representations. In the case of the hard drive, a web shop may first offer some criteria the agent can choose; only then he will see a list of results. To exploit this, the agent must adapt to these different layers. For example, the agent may first filter all possible results and then look at the resulting list. With the right sequential composition of subprocedures adapted to the environment, the agent can find his choices equivalent to full maximization.

One may wonder, if we focus on rationalizable behavior, what is the point of considering the environment? After all, choices are in line with maximization. We have two answers. First, even if we can rationalize choices and ignore the environment and its constraints, it does not mean we have a good description of what people choose. Considering the environment helps to characterize preferences. For instance, if we know alternatives are ordered according to some criterion and the agent uses this criterion, we can pin down his preferences more precisely. We cannot do this without incorporating the environment. Moreover, looking in the other direction, if we observe how people proceed when buying, say in an online environment, what may appear as not rational at first sight, can be part of choosing rationally in the bigger picture. Again, choosing the first element of a list of alternatives is such a case. In itself it does not make too much sense through the lens of rational choice but once seen as being part of a sequence of procedures, it can make a lot of sense. Secondly, while we are focused on rational choice, the environment constrains in general how an agent can proceed. This is true independently of the underlying decision criterion.

1.1 Related work

Our work is related to two strands of the literature: sequential search and bounded rationality. The idea that the environment is important in shaping the choice process is of course not novel but goes back at least to Herbert Simon [15]. As we model the decision-making process, an elementary part of it is search. When facing a list of alternatives, a decision-maker inspecting the elements step by step is searching. Sequential search processes as the latter example were already studied by Simon [14]. More recent (and more general) investigations of search behavior and its rationalizability are [2, 5]. Our work is different as we model the environment and how it interacts with search.

Regarding the literature on bounded rationality, our paper is most closely related to two papers by Ariel Rubinstein and Yuval Salant [10, 12]. In both papers, a decision problem is enriched by additional structure. Rubinstein and Salant ask, if one ignored this extra structure but focused only on the choice function, respectively choice correspondence which is induced by a choice procedure operating on the richer structure, under which conditions can choices be rationalized. They identify conditions which guarantee rational choices. They also show that ignoring the extra structure is problematic from an economic perspective. Even if choices are rationalizable, the derived preference relation carries no meaning whereas the procedure on the richer structure has a clear economic interpretation.

We build on their work. We also ask when an induced choice function/correspondence can be rationalized given a representation and a procedure, we adopt some of their axioms to our setting and show that they also guarantee rationalizability in our framework, and, lastly, we also provide evidence that ignoring the richer structure is problematic if one wants to provide an economic interpretation.

With all these similarities what is the difference then? In [10], the extra structure imposed is that alternatives are organized as lists. In our framework their version of lists is a special case, considering the particular context-free grammar that described lists. Yet, we also consider other representations such as lists of lists, trees, etc. Therefore, the results we derive on rationalizability are not limited to one particular representation. This is also important as there are several papers following the idea of [10] and consider various other special cases [16, 6, 3]. While we generalize the results, our approach is also more “expressive”. In [10] a list is a global property of the set of alternatives. In our setting, we consider the local constraint under which an agent can access the alternatives. Maybe he can only see the first element, maybe he can see the whole list at once. These different constraints arise from different representations of lists, while they are the same in [10]. In that sense, we are more specific with respect to the modeling of the constraints while at the same time our results are more general as they are based on the properties of the representation and not their concrete realization.

In [12], the extra structure imposed is a “frame” turning a classical choice problem into an “extended choice problem”. A frame capture different things such as the representation of alternatives, e.g. in the form of a list, the timing of decisions, but also internal restrictions of the agents. Their frames capture broader phenomena than what we focus on in our paper. We focus on the representation of alternatives and how they can be accessed. Their “extended choice function” which operates on the extended choice problem is analogous to the classical choice function which maps input (budget) into output (choice). In contrast, a choice procedure in our setting models the choice process locally, i.e. how the agent proceeds step by step. Choice processes so can be modeled recursively, operating on also recursively defined representations. Our formulation also caters for compositionality: representation can be formed from sub-representations, and choice procedures can be constructed from sub-procedures. We believe compositionality is key when we want to describe actual decision-making. If you by a hard drive and you use filter and sort functionality to reduce the search results, then look at the first two pages of results to choose one element, then your choice process is composed of different subprocesses. What is more, the interaction is relevant for the economic interpretation. If filtering and sorting capture your preferences, then your procedure on the list of search results can be simple: You just choose the first alternative. As [12] showed, for making sense of decisions, considering extra structure is key. We add to this that the composition of subprocesses is key to make sense of the overall meaning.

The paper is organized as follows. In Section 2, we introduce the notion of a represented decision problem, a decision procedure, and give examples of represented problems and procedures. We provide several properties of procedures which enable us to characterize classes of procedures in Section 3. Using

these properties, in Section 4 we then characterize under which conditions procedures lead to choices that can be rationalized. Lastly, in Section 5 we discuss how one can attach different meanings to the representation and how composites of procedures fit into our framework.

2 Represented Decision Problems and Decision Procedures

Choice theory is concerned with decision problems and choice functions. A decision problem is modeled as a subset A of a set of alternatives X , i.e. $A \subseteq X$ and the choice function $c: \mathcal{P}(X) \rightarrow X$ as a function operating on the decision problem, using $\mathcal{P}(X)$ to denote the powerset of X . This characterizes the behavior of an agent.

The representation of alternatives as a *set* abstracts from the restrictions of the environment, and how the alternatives are actually presented to the decision-maker. However, in practice agents do not see the set of alternatives (at once). For instance, some online retailers only offer a single alternative at a time. To access more options agents have to navigate to the next item. Or some online shops will first ask you specific questions one at a time, and depending on the answers given they will offer you a single alternative. Our first task is to model these “access restrictions”, by formalizing how decision problems are represented.

2.1 Representing decision problems with algebraic datatypes

We think of represented decision problems as words in a given language, which we formalize using context-free grammars. We will use an extension of context-free grammars known in Computer Science as *algebraic datatypes* (ADT). Algebraic datatypes extend context-free grammars in two crucial ways. First, languages are described parametrically on the set X . Applied to our first example, we have

$$\text{Item } X \Rightarrow X$$

$$\text{Page } X \Rightarrow [\text{Item}_1 X, \dots, \text{Item}_{10} X]$$

$$\text{Result } X \Rightarrow \text{Page } X, \text{Result } X$$

so that we can uniformly talk about a search for books (Result Book) or for hard drives (Result Harddrive), where Book and Harddrive here the set of all books or hard drives, respectively.

As a second extension we attach to each rule of the grammar a *constructor*. The constructor acts as a tag. Given a problem representation, it allows us to reconstruct how the representation was actually generated. Again, using our introductory example, we have three constructors, one for each of the rules

$$\text{Item } X \Rightarrow I X$$

$$\text{Page } X \Rightarrow P (\text{Item}_1 X) \dots (\text{Item}_{10} X)$$

$$\text{Result } X \Rightarrow R_1 (\text{Page } X) \mid R_2 (\text{Page } X) (\text{Result } X)$$

So, a Result consisting of two pages has the form

$$R_2 (P (I x_{1,1}) \dots (I x_{1,10})) (R_1 (P (I x_{2,1}) \dots (I x_{2,10})))$$

The general definition of ADTs is as follows:

Definition 2.1 (Algebraic Datatype, ADT) *Let X stand for an arbitrary set. An algebraic datatype TX is a datatype defined via a generalised context-free grammar of the form*

$$TX \Rightarrow C_1 t_{1,1} \dots t_{1,n_1} \mid \dots \mid C_k t_{k,1} \dots t_{k,n_k}$$

where C_i , for $1 \leq i \leq k$ are called the data constructors, $n_i \geq 1$ is the arity of the i -th constructor, and $t_{i,j}$ is either in X or TX .

Remark 2.2 Algebraic datatypes play an important role in the foundations of functional programming. As choice procedures can be expressed in this language, we get an implementation in code for free. To illustrate the relation between theory and code, we include Haskell code for the representations and procedures used in this paper (see the Appendix).¹

We call a decision problem over X represented by an algebraic datatype TX as a *represented decision problem*. Given a set of alternatives X , an ADT has a finite list of constructors through which the alternatives can be accessed. To be concrete, let us consider a few more examples:

- **List**. For any set X we can consider the datatype **List** X of non-empty lists over X defined from the singleton list **Sing** x and a data constructor **Cons** that prepends a new element to a given list.

$$\mathbf{List} X \Rightarrow \mathbf{Sing} X \mid \mathbf{Cons} X (\mathbf{List} X)$$

This definition tells us that a **List** X is defined by two types of constructors (the right hand side of the equation): A list is either a single element “**Sing** x ” or (represented by “**|**”) an element “ x ” being glued at the front of an existing list, the second part “**Cons** x (**List** x)”. Note that these are recursively defined datatypes; and **Sing** and **Cons** should be viewed as mappings of type²

$$\mathbf{Sing} : X \rightarrow \mathbf{List} X \quad \mathbf{Cons} : X \rightarrow \mathbf{List} X \rightarrow \mathbf{List} X$$

Also note, for now, we assume “ x ” to be an element of a given set of alternatives. But in principle, “ x ” can have more structure as well. We will come back to this.

The constructors are important as they define how an agent can access the list. Under **List** an agent can only access the single element at the head of the list, or the “tail” of the list. If he wishes to inspect the next elements, he can do so only step by step. This sequential feature of representations is ubiquitous in online search. That only one element can be inspected at a time is of course an extreme case. It is straightforward to consider extensions to multiple alternatives being accessible at once.

An example of an element in the datatype **List** X is **Cons** x (**Cons** y (**Sing** z)), where x, y, z are elements of X . We will abbreviate a list such as **Cons** x (**Cons** y (**Sing** z)) as $[x, y, z]$.

- **List**₂. We can also consider a representation of lists where list-concatenation is a primitive constructor:

$$\mathbf{List}_2 X \Rightarrow \mathbf{Sing}_2 X \mid \mathbf{Cat} (\mathbf{List}_2 X) (\mathbf{List}_2 X)$$

Different representations of lists have consequences for how an agent can choose. In contrast to **List** where the agent has access to alternatives only step by step, in the case of **List**₂ an agent has access to the complete list at once and can “break” any list into two sublists. An example for such a representation is a menu ordered by price. The agent may look for a price threshold up to which he is willing to pay and then consider two sublists below and above the threshold.

- **Tree**. Consider the datatype of non-empty trees **Tree** X defined from singleton node **Leaf** x and a parent node with some fixed number of sub-trees:

$$\mathbf{Tree} X \Rightarrow \mathbf{Leaf} X \mid \mathbf{Node} (\mathbf{Tree} X) \dots (\mathbf{Tree} X)$$

For instance, if $X = \{a, b, c\}$ and the arity of **Node** is three, then some of the elements of **Tree** X are **Leaf** a and

$$\mathbf{Node} (\mathbf{Leaf} a) (\mathbf{Leaf} b) (\mathbf{Node} (\mathbf{Leaf} a) (\mathbf{Leaf} b) (\mathbf{Leaf} c))$$

An example for this representation would be an online shop where a customer navigates a website structured in a tree-like fashion.

¹Haskell is a modern functional programming language (see <https://www.haskell.org/>) for information.

²Note, this notation is known as “currying”. The evaluation of a function with several arguments can be equivalently translated into an evaluation of a sequence of functions with a single argument each.

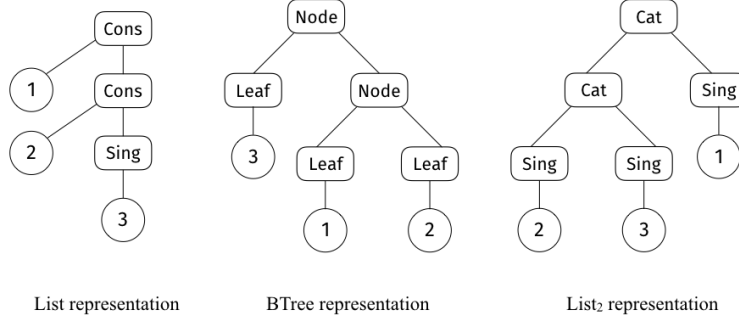


Figure 1: Possible different representations of the set $\{1, 2, 3\}$

- **LTree**. We can also consider trees whose nodes are labeled. In its simplest form this can be achieved by allowing different types of node constructors. For simplicity here we assume that all nodes have arity 2 (two sub-trees):

$$\mathbf{LTree} X \Rightarrow \mathbf{LLeaf} X \mid \mathbf{Node}_1 (\mathbf{LTree} X) (\mathbf{LTree} X) \mid \dots \mid \mathbf{Node}_2 (\mathbf{LTree} X) (\mathbf{LTree} X)$$

We already mentioned that trees play an important role in representing restrictions agent face in online shopping. The labelling adds more economic content by attaching meaning to the nodes. For instance, labels can represent properties of goods such as “red” vs. “white”. An agent who filters alternatives based on such properties then will navigate the tree accordingly.

We note that ADTs are essentially structured labelled trees. For instance, Figure 1 shows the tree representation of the set $\{1, 2, 3\}$ as either a **List**, **List₂** or **Tree**. However, although a labelled tree can be arranged in any possible way, for a labelled tree to be a valid tree representation of an algebraic datatype it has to abide by the rules of the ADT grammar. For instance, Figure 2 shows two labelled trees which do not correspond to any of the datatypes above. The first tree uses the right constructors (label) as in **List**, but in the wrong places, as **Sing** only takes one argument, and **Cons** should take two arguments. In the second case we have a labelled tree which uses constructors from both **Tree** and **List₂**.

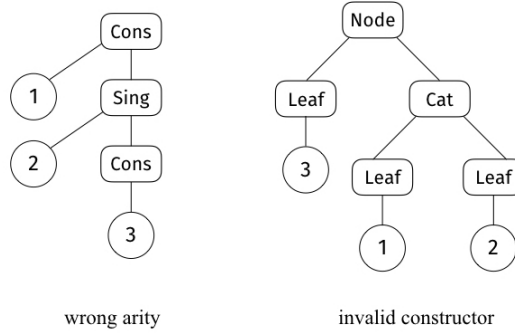


Figure 2: Some labelled trees which do not correspond to the ADTs above

Remark 2.3 (Recursive ADTs) When $t_{i,j} = TX$ for some i, j we say that the datatype TX is recursive, as in all the examples above. In general, recursively defined ADTs allow for infinite data structures, e.g. infinite lists or trees – a feature which plays an important role in their use in programming [7]. In this paper, however, we only consider finite elements of TX . This means that when navigating a finite element

$a: TX$ we must always reach a constructor C_j where all arguments $a_1 \dots a_{n_j}$ are elements of X (and not elements of TX). In the context of representing choices, we obviously want to restrict our attention to finite representations.

Algebraic datatypes are very expressive. The two representations of lists introduced above illustrate this. As we will see later, how agents can choose on different constructors varies. This has consequences for the rationalizability of procedures. For our purpose in this paper, ADTs are crucial because they formalize the common structure which is shared by *different* constraints of the environment. So, instead of providing a particular representation for lists, and another for lists of lists, and yet another for trees etc. for which a theory will be limited, ADT allow us to introduce conditions for which the theory *generally* holds.

So far, we only have described how representations of particular decision problems are modeled but have been silent on whether this is possible for every decision problem. To that end consider the next property.

Definition 2.4 (T-representation) Let 2^X denote the set of non-empty finite subsets of X . A mapping $r: 2^X \rightarrow TX$ will be called a *T-representation* if $A = \llbracket r(A) \rrbracket$, for any $A \subseteq X$.

A *T-representation* specifies how a classical problem gets represented as an element of TX . We will generally assume in this paper that the datatypes TX admit a *T-representation*. This is required in order to ensure that any decision problem $A \subseteq X$ can be represented by an element $a \in TX$, i.e. $\llbracket a \rrbracket = A$. All of the standard algebraic datatypes such as lists and trees admit a representation. Some peculiar examples of ADTs which do not allow for representation of arbitrary decision problems are

- *flat* ADTs, i.e. ADTs that do not make use of recursion, e.g.

$$TX \Rightarrow C_1 X \mid C_2 X X$$

which can only represent decision problems of size 1 or 2;

- *non-productive* ADTs, i.e. ADTs that do not have a non-recursive constructor, e.g.

$$TX \Rightarrow C_1 (TX) (TX) \mid C_2 (TX)$$

which only generates infinite trees without any leaf (value) node – and hence represents the empty set;

It is easy to argue that these are the only two cases where an ADT will not be able to represent decision problems of arbitrary size.

Proposition 2.5 An ADT TX is able to represent decision problem of arbitrary size if and only if it is *productive and non-flat*.

Proof. In one direction, assume the ADT is productive and non-flat, i.e. it has both recursive and productive constructors, e.g.

$$TX \Rightarrow \dots C_i (TX) (TX) \mid \dots \mid C_j X \dots$$

Then using C_i and C_j one can build labelled trees with an arbitrary number of leaf (value) nodes, which can then be populated with values from an arbitrary decision problem. For the other direction, the example above shows that an ADT which is able to represent problems of arbitrary size must be productive and non-flat, noting that ADTs are only allowed to define finitely many constructors. \square

2.2 The extension mapping on represented decision problems

For a given represented decision problem, we can extract the elements it contains to get the decision problem.

Given a represented decision problem $a \in TX$ we will denote its *extension* by $\llbracket a \rrbracket \subseteq X$. This mapping can be defined by induction on the definition of TX

$$\llbracket C_i r_1 \dots r_{n_i} \rrbracket = \bigcup_{1 \leq j \leq n_i} \llbracket r_j \rrbracket$$

where $\llbracket r_j \rrbracket = \{x\}$ if $r_j = x$. This is well-defined on well-founded (i.e. finite) represented problems $a \in TX$. So the mapping $\llbracket \cdot \rrbracket$ translates represented decision problems into their set extension, e.g.

$$\llbracket [x, x, y] \rrbracket = \llbracket [y, x] \rrbracket = \{x, y\}$$

Hence, the extension mapping “extracts” the classical decision problem from a represented decision problem. Therefore, a represented decision problem is always an enriched classical decision problem. And the latter can always be retrieved by abstracting away the representation.

Definition 2.6 (Equivalent represented decision problems) *We say that two represented decision problems $a_0, a_1 \in TX$ are equivalent, written $a_0 \sim a_1$, if $\llbracket a_0 \rrbracket = \llbracket a_1 \rrbracket$.*

So, formally we have that $\llbracket [x, y] \rrbracket$ is equivalent to $\llbracket [y, x, x] \rrbracket$, i.e. $[x, y] \sim [y, x, x]$.

Let us conclude this section on representations by defining a special class of algebraic datatypes which we will use in Proposition 3.5:

Definition 2.7 *An algebraic datatype TX is called substitutable if all elements $x \in \llbracket a \rrbracket$ must appear in $a \in TX$ under a unary constructor.*

A substitutable representation allows a subproblem to be replaced by a single element. To anticipate our discussions of procedures, a subproblem could be replaced by an element that a procedure chooses on that particular subproblem. Thus, an agent may decide on the subproblems first and then consider the “reduced” representation of these choices to come to an overall decision.

For substitutable datatypes we can define a substitution operation $a[b/x]$ which locally replaces an element of x by a new (represented) sub-problem $b \in TX$, as follows. For unary constructors we define:

$$\begin{aligned} (C_i y)[b/x] &= C_i y, \text{ when } y \neq x \\ (C_i x)[b/x] &= b \end{aligned}$$

whereas for non-unary ones we let:

$$(C_j a_1 \dots a_{n_j})[b/x] = C_j (a_1[b/x]) \dots (a_{n_j}[b/x])$$

Tree and **List**₂ are examples of substitutable datatypes, whereas **List** is not substitutable. For instance, consider the datatype **Tree** X where **Tree** is defined in Section 2.1. We have that

- $a = \text{Node}(\text{Leaf } x_1) (\text{Leaf } x_2) (\text{Leaf } x_3)$ and
- $b = \text{Node}(\text{Leaf } x_4) (\text{Leaf } x_5) (\text{Leaf } x_6)$

are elements of **Tree** \mathbb{N} . Then $a[b/x_2]$ is the element

$$\text{Node}(\text{Leaf } x_1) (\text{Node}(\text{Leaf } x_4) (\text{Leaf } x_5) (\text{Leaf } x_6)) (\text{Leaf } x_3)$$

i.e. we have replaced the value x_2 in tree a by the sub-tree b .

2.3 Decision procedures

Decision procedures describe how an agent computes a choice given a represented decision problem.

Definition 2.8 (Decision procedure) *We will call a decision procedure a function that selects an element of X given a represented decision problem. Formally, a decision procedure is a mapping $P: TX \rightarrow X$ such that $P(a) \in \llbracket a \rrbracket$, for all $a: TX$.*

Decision procedures are defined by a case-distinction on each of the constructors of TX , i.e. for each constructor we need to say how the procedure will operate, and what will be the next step in the decision process be. Thus, the constructors influence how an agent can choose given a representation. Let us illustrate this point with our representations of **List**. Suppose the agent always chooses the first element of the list. His procedure can be described as

$$F: \text{List } X \rightarrow X$$

$$F(\text{Sing } x) = x$$

$$F(\text{Cons } x \ xs) = x$$

Given $(\text{Sing } x)$, a single element list, the procedure chooses this single element. Given $(\text{Cons } x \ xs)$, a list containing a head and an yet undetermined tail, the procedure chooses the head. Since any element of **List** is either one of the two forms, the procedure is a well defined function.

Consider another simple example. A **Tree**-decision procedure that always selects the left-most element on the tree can be defined as

$$L(\text{Leaf } x) = x$$

$$L(\text{Node } t_1 \ t_2 \ t_3) = L(t_1)$$

This procedure defines for each constructor what the next step is. Note, this procedure is recursive. In the second case, $(\text{Node } t_1 \ t_2 \ t_3)$, the procedure will call itself on the left branch of the given tree. For this procedure one can therefore describe how the global choice process evolves from the local description. In the case of the procedure before, choosing the first element of a list, local and global are the same as there is just one step.

2.4 Relation to choice functions

How are procedures related to classical choice functions? We begin with taking a choice function as given. We can use the extension mapping on decision problems to lift standard choice functions $c: 2^X \rightarrow X$ into decision procedures $\llbracket c \rrbracket: TX \rightarrow X$ by simply composing the two

$$\llbracket c \rrbracket(a) = c(\llbracket a \rrbracket)$$

or more concisely $\llbracket c \rrbracket = c \circ \llbracket \cdot \rrbracket$. We will call $\llbracket c \rrbracket$ the *canonical decision procedure* associated with the choice function $c: 2^X \rightarrow X$.

Note, analogous to a classical decision problem being the stripped-down version of a represented decision problem, a classical choice function is a stripped-down version of a decision procedure. It is always possible to implement a choice function as a decision procedure by the lifting defined above.

In the other direction, for a particular T -representation, we can define a choice function that is implied by a given decision procedure. Formally,

Definition 2.9 (Choice function from decision procedure) *Given a decision procedure $P: TX \rightarrow X$ and a T -representation, define the choice function $\{\{P\}\}_r: 2^X \rightarrow X$ as*

$$\{\{P\}\}_r(A) = P(r(A))$$

2.5 Discussion

What we have introduced so far is the bare minimum to model constraints imposed by the choice environment. The only added piece of content compared to a classical decision problem is that it is represented according to some construction. Yet, actual representations may provide richer information. Suppose you want to buy a hard drive. You only care about capacity. You receive a list of search results which you can order according to capacity. Then you can just choose the first element. So far we have not touched such restrictions and we have also not discussed how we handle attributes of goods – such as capacity.

We will first stick to the bare minimum and develop our theory based on this basic notion of representation. Why do we start with the possibly less realistic basic representation instead of directly moving to

the economically more relevant scenarios? First, the theory we introduce for the simpler cases seamlessly extends to the more complicated cases. Secondly, and more importantly, there is a crucial modeling aspect: What information does the agent have access to? And moreover, does an outsider, who sees the choices an agent makes, also know this information? Consider the list of hard drives ordered according to capacity. If the agent is aware of this information, his procedure can be different compared to the situation where he does not have this information. If he knows and his preferences are about capacity, he can just choose the first element. Next, consider an outside observer, who wants to make sense of an agent's choices. Suppose, he observes that the agent always chooses the first element but does not know that the list is ordered according to capacity. If the observer asks whether and how these choices are rationalizable, this piece of information is important. And therefore, this question depends on the representation and what information an outsider has access to. Now, what is best to be assumed is not set in stone. It depends on the situation.

The minimum assumption we make here is that an outside observer knows the type of representation the agent is using. We assume this to be true in all cases. We start from this minimum, and the theory we introduce will first operate on this assumption. Then, in Section 5 we will gradually enrich the representation. First, we will introduce *guarantees*, that is, constraints on the representation, like the list of hard drives having always a fixed order. We will also consider cases where economic meaning is attached to the representation, like the list of hard drives being ordered according to capacity. We will show that these richer environments affect the possibility how to rationalize choices as well as influence the economic interpretation of procedures.

Before we continue, let us also state two simplifications we make in this paper. First, the externally designed representation is the basis of the internal procedure by the agent. That is, if the alternatives are represented as a list, the agent's procedure will operate on a list. This is clearly a simplification. It is possible that the agent's information processing transforms the representation before he makes a decision. We consider the simpler case a natural starting point for considering information-processing in general. Moreover, as argued before, given today's technology, representations are designed and at least for some environments observable. It is much harder (if possible at all) to describe internal transformations from the vantage point of an external observer. Lastly, from a theoretical point, there is no change needed to accommodate internal transformations. The only missing piece is a transformation of a datatype; procedures operating on the datatype can be reused; and as procedures can be composed (more on that in Section 5), the transformation operation would be a further element in a sequence of operations.

A second restriction concerns the formalization of behavior. In our setup, given an internal procedure we ask when the resulting behavior can be mapped into a choice function. This is not the only possibility. Indeed one could interpret the extended choice function in [12] as a richer behavioral model and the same for the extended choice function in [5]. We are sympathetic to the idea that the behavioral model should be broadened, in the end information technology has transformed what we can observe and measure. Still, by introducing representations we already made a significant change. To understand what this change entails, we keep the behavior the same.

2.6 Examples of decision procedures

Let us consider some examples of decision procedures. These examples serve two purposes. First, they provide some illustrations of how to use ADT's and procedures to model decision processes. Secondly, we will use some examples for the proofs in the next section. These examples are more of a technical nature. They provide evidence that procedures with certain properties do exist *in principle*. We do not claim that such examples are realistic descriptions of actual procedures – at least not when viewed in isolation. We will come back to this point in Section 5.

Appendix 7.2 provides the Haskell code for the examples. Why do we do this? First, representation in code makes it easy to explore the procedures. You can run them and change them quickly. Secondly, for economic theorists, it might be interesting to realize how close theory and code are. Modelling in our framework almost automatically means having running code available. Thirdly, being able to compile the procedures is the best proof that the procedures are formally correct and well-defined. On paper you can trick yourself but you cannot trick the Haskell compiler.

Example 2.10 (Maximizing) Consider a procedure $P: TX \rightarrow X$ that recursively goes through a represented decision problem $a \in TX$ to find the maximal element according to some predetermined strict preference relation (X, \succ) , i.e. a complete, transitive and anti-symmetric relation. Such procedure can be recursively defined as: For each constructor C_i of TX (of arity n_i)

$$P(C_i a_1 \dots a_{n_i}) = \max_{\succ} \{a_j^* \mid j \in \{1, \dots, n_i\}\}$$

where $a_j^* = P(a_j)$ if $a_j \in TX$ and $a_j^* = a_j$ if $a_j \in X$.

Note that this procedure can operate on different representations.

Example 2.11 (List Satisficing) Consider the satisficing **List**-decision procedure whereby an agent has an evaluation function $u: X \rightarrow \mathbb{R}$ and a satisficing³ threshold $u^* \in \mathbb{R}$. The agent chooses the first element of the list that is above the given threshold. If there is none, he chooses the last element. This procedure can be recursively defined as:

$$P(\text{Sing } x) = x$$

$$P(\text{Cons } x a) = \text{if } u(x) \geq u^* \text{ then } x \text{ else } P(a)$$

Example 2.12 (List₂ Satisficing) Consider the satisficing **List₂**-decision procedure whereby an agent has an evaluation function $u: X \rightarrow \mathbb{R}$ and a satisficing threshold $u^* \in \mathbb{R}$. The agent chooses the left-most element of the list that is above the given threshold. If there is none, he chooses the last element. This procedure can be recursively defined as:

$$P(\text{Sing}_2 x) = x$$

$$P(\text{Cat } a_l a_r) = \text{if } u(P(a_l)) \geq u^* \text{ then } P(a_l) \text{ else } P(a_r)$$

Example 2.13 (Conditional List₂ satisficing) The set of alternatives X is partitioned into two sets, K (known, or default) and N (new). Given a problem $a \in \mathbf{List}_2 X$, if more than half of the elements in $\llbracket a \rrbracket$ are known elements, choose the first element of the list which is in K . Otherwise choose the maximal element of the list according to some predetermined order (X, \succ) as in Example 2.10.

Example 2.14 (Choose default on large problems) Assume the agent has a default choice $x \in X$ in mind and a size threshold N . Consider the procedure that given $a \in TX$ returns x when $x \in \llbracket a \rrbracket$ and $|\llbracket a \rrbracket| > N$, or else returns the maximal element in $\llbracket a \rrbracket$ according to some predetermined order (X, \succ) as in Example 2.10.

Example 2.15 (First option on list) Consider procedure *first* from Section 2.3) which always returns the first element of a given list.

We can also implement this procedure on **List₂**

$$P_2(\text{Sing}_2 x) = x$$

$$P_2(\text{Cat } a_l a_r) = P(a_l)$$

Example 2.16 (Second option on list) Consider also a variant of the above, where the agent always chooses the second element of the list, or the first element of a list on singleton lists:

$$P(\text{Sing } x) = x$$

$$P(\text{Cons } x(\text{Cons } y ys)) = y$$

Again, we can also implement the same procedure on **List₂**:

$$P_2(\text{Sing}_2 x) = x$$

³This example is based on [9, p.32]. Satisficing goes back to [14].

$$P_2(\text{Cat}(\text{Sing}_2 x) a_r) = P_2(a_r)$$

$$P_2(\text{Cat} a_l a_r) = P_2(a_l)$$

again we can implement it on list 2

Example 2.17 (Larger sub-problem bias) Consider an agent choosing on trees $a \in \text{Tree } X$, but this time suppose she also has a size threshold $N \geq 1$. The agent always recursively chooses to work on the biggest of the three sub-trees, and maximises when reaching a sub-problem of size smaller or equal to N :

$$P(a) = \max_{\succ} \llbracket a \rrbracket, \text{ if } |\llbracket a \rrbracket| \leq N \quad (\text{as in Example 2.10})$$

$$P(\text{Node } t_1 t_2 t_3) = \text{if } |\llbracket t_1 \rrbracket| > \max(|\llbracket t_2 \rrbracket|, |\llbracket t_3 \rrbracket|) \text{ then } P(t_1) \text{ elseif } |\llbracket t_2 \rrbracket| > |\max(|\llbracket t_1 \rrbracket|, |\llbracket t_3 \rrbracket|)| \text{ then } P(t_2) \text{ else } P(t_3)$$

Example 2.18 (Smaller sub-problem bias) Consider a variation of the example above where the smaller of the two sub-problems is preferred:

$$P(a) = \max_{\succ} \llbracket a \rrbracket, \text{ if } |\llbracket a \rrbracket| \leq N \quad (\text{as in Example 2.10})$$

$$P(\text{Node } t_1 t_2 t_3) = \text{if } |\llbracket t_1 \rrbracket| < \min(|\llbracket t_2 \rrbracket|, |\llbracket t_3 \rrbracket|) \text{ then } P(t_1) \text{ elseif } |\llbracket t_2 \rrbracket| < |\min(|\llbracket t_1 \rrbracket|, |\llbracket t_3 \rrbracket|)| \text{ then } P(t_2) \text{ else } P(t_3)$$

Example 2.19 (Avoiding undesirable element) Consider a procedure for selecting an element from a list of alternatives $a \in \text{List}_2 X$. The procedure will recursively operate on the left sub-tree and take that choice, except when that choice is an undesirable element $x^* \in X$ and the number of choices is large (bigger than some given threshold $N \geq 1$), in which case it disregards the left choice and works recursively on the right sub-tree. The element x^* and the threshold N are fixed exogenously. The procedure can be defined as:

$$P(\text{Sing}_2 x) = x$$

$$P(\text{Cat } a_l a_r) = P(a_l), \text{ if } P(a_l) \neq x^* \text{ or } |\llbracket a_l \rrbracket \cup \llbracket a_r \rrbracket| \leq N$$

$$P(\text{Cat } a_l a_r) = P(a_r), \text{ if } P(a_l) = x^* \text{ and } |\llbracket a_l \rrbracket \cup \llbracket a_r \rrbracket| > N$$

3 Properties of Decision Procedures

We want to characterize procedures which allow a rationalization. To that end we introduce several properties. Let us begin by defining two properties that are closely related to properties in classical choice theory.

3.1 Classical consistency

Definition 3.1 ($\alpha\mathbf{E}$) A procedure $P: TX \rightarrow X$ is said to satisfy the α -extended property if

$$\forall a \in TX (P(a) \in B \subseteq \llbracket a \rrbracket \Rightarrow \exists b \in TX (\llbracket b \rrbracket = B \wedge P(b) = P(a)))$$

Thus, if a procedure P fulfills $\alpha\mathbf{E}$, then if x is chosen under the represented problem a , and one considers a subset B of the elements contained in the represented problem a , then there exists a representation of the set B for which x is chosen by procedure P .

Definition 3.2 ($\gamma^+\mathbf{E}$) A procedure $P: TX \rightarrow X$ is said to satisfy the property γ^+ -extended if

$$\forall a, b \in TX (P(b) \in \llbracket a \rrbracket \Rightarrow \exists d \in TX (\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket \wedge P(d) = P(a)))$$

Thus, if a procedure P fulfills $\gamma^+\mathbf{E}$, then if x is chosen under the represented problem a , y is chosen under the represented problem b , and y is an element of the set of alternatives represented in a , then there exists a representation of the union of the sets of alternatives represented in a and b .

These two properties follow closely their classical counterparts in choice theory: they impose consistency of choices if a problem gets reduced or if two problems get combined. We adapt them to fit into our language. The motivation to include them in the first place goes back to [12]. They introduce adaptations of α and γ in their language of extended choice functions to characterize which extended choice functions can be rationalized. In the next section, we see that they play a similar albeit not the same role in our language.

3.2 Procedural consistency

Both, $\alpha\mathbf{E}$ and $\gamma^+\mathbf{E}$, only guarantee that a representation does exist which fulfills the consistency of choices. They do not provide details how such representations ought to be constructed given the initial represented problem. In the next section, we will ask when it is possible to provide such a concrete construction. Moreover, we will impose consistency of choices at each step of this transformation so that in the end $\alpha\mathbf{E}$, respectively $\gamma^+\mathbf{E}$, is fulfilled.

To this end we introduce the following two properties:

- A decision procedure is *strongly inductive* (**SIND**), if the choice of a given problem is always the same as the choice in one of the immediate sub-problems, and this choice does not depend on the structure of the sub-problems, i.e.

$$\forall C_i \forall x_1, \dots, x_{n_i} \in X \exists k \in \{1, \dots, n_i\} \forall a_1, \dots, a_{n_i} \left(\bigwedge_{j=1}^{n_i} x_j = a_j^* \rightarrow x_k = P(C_i a_1 \dots a_{n_i}) \right)$$

where $a_j^* = P(a_j)$ if $a_j \in TX$ and $a_j^* = a_j$ if $a_j \in X$.

- Assume TX is substitutable. A procedure $P: TX \rightarrow X$ is said to satisfy *T-independence of irrelevant alternatives* (**TIIA**) if given any problem $a = C_i \dots a' \dots$ such that $P(a) = x$ and $P(a') = y$ and $x \neq y$, then there exists a b such that $\llbracket b \rrbracket = (\llbracket a' \rrbracket \setminus \{y\}) \cup \{x\}$ and $P(C_i \dots b \dots) = x$.
Intuition: If a y is chosen on an immediate sub-problem, but not on the overall problem, we should be able to replace y by x on that immediate sub-problem without affecting the overall choice.

The last property is an adaption of the classical axiom to our framework. Note that this property imposes consistency of choices when we reduce a given problem a by substituting one element with the chosen element on one of its subproblems. Later we will use this property to inductively reduce a given problem with elements $\llbracket a \rrbracket$ to $\llbracket b \rrbracket$. Note that such a reduction does not change the shape of the representation. Consider a given tree where successively elements get replaced by an x . The overall number of nodes is still the same.

Strongly inductive captures the notion of “divide-and-conquer”. An agent decides by dividing the overall problem in subproblems, and then choosing on the subproblems. The overall choice is then determined by the choices on the subproblems. The notion of divide-and-conquer was introduced by [8] under the name path-independence. We also note that **SIND** is related to *partition independence*, a property introduced in [10] as their interpretation of Plott’s path independence property. It is defined in the context of their representation of lists. We note here that partition independence is a considerably stronger notion. In combination with their definition of lists, in [10] agents can break up a list at arbitrary positions. Partition independence implies that procedures are insensitive to where the list has been separated. In our framework, agents can break up the problem only in so far as the constructors do allow this. And in general, procedures will react differently depending on which representation is used. We can strengthen representations and procedures operating on them, but this is not our default. We come back to this and related issues in Section 5.

Which formulation is more reasonable from a modeling perspective? We think there are situations where the agent observes the full list and can split the list wherever he likes so that the condition introduced in [10] is a faithful representation of the agent’s perspective. On the other hand, there are situations where the representation is not fully under the agent’s control and he can only access the representations piece-wise. This is in particular true for online shopping, online experiments, surveys etc. where agents do not have access to the whole representation right from the beginning. For that modeling the constructors explicitly is a better choice.

The following proposition clarifies the relation between **TIIA** and **SIND**.⁴

Proposition 3.3 *Assume TX is substitutable. **SIND** and **TIIA** are incomparable, i.e.*

- (i) *There is a procedure $P: TX \rightarrow X$ which is **SIND** but not **TIIA**.*

⁴This is a further difference to [10] where both conditions are equivalent for lists without duplications and their form of list independence are equivalent under no duplication and list independence is implied by partition independence under duplications.

(ii) There is a procedure $P: TX \rightarrow X$ which is **TIIA** but not **SIND**.

Proof. (i) Let $X = \{x, y, z\}$. Assume the circular binary relation on X defined as $x \succ y$ and $y \succ z$ and $z \succ x$. Finally, define the procedure $P: \text{Tree } X \rightarrow X$ defined as

$$P(\text{Leaf } x) = x$$

$$P(\text{Node } t_1 t_2 t_3) = \max_{\succ} \{P(t_1), P(t_2), P(t_3)\}$$

This procedure is clearly **SIND**. We show that it is not **TIIA**. Consider the tree

$$a = \text{Node}(\text{Leaf } x)(\text{Leaf } x)(\text{Node}(\text{Leaf } y)(\text{Leaf } z)(\text{Leaf } z))$$

We have that $P(a) = x$. But as soon as we replace y by x on the right sub-tree the overall choice will be z .

(ii) Consider Example 2.17 (**Larger sub-problem bias**). This procedure is not **SIND** since the decision relies on looking at the actual size of the sub-problems. To see that it satisfies **TIIA**, note that $\llbracket b[x/y] \rrbracket \leq \llbracket b \rrbracket$. Hence, since the procedure always chooses to go to the biggest sub-problem, this substitution will not affect the operation of the procedure. \square

3.3 Extensional and intensional procedures

We end our characterization of procedures by describing two fundamentally different classes of procedures.

- A decision procedure is *extensional* (**EXT**) if it does not make use of the representation of the problem when making the decision, i.e. it produces the same outcome on different representations of the same decision problem:

$$\forall a, b \in TX (a \sim b \Rightarrow Pa = Pb)$$

Recall that $a \sim b$ is defined as $\llbracket a \rrbracket = \llbracket b \rrbracket$, i.e. a and b are representations of the same set.

- A decision procedure is *intensional* (**INT**) if it does not make use of the actual values of the represented decision problem, but chooses simply based on the representation structure:

$$\forall a \in TX \forall x, y \in X (P(a) = x \Rightarrow P(a[y/x]) = y)$$

where $a[y/x]$ denotes the action of replacing all occurrences of x in a by y .

These two properties capture extreme classes of decision procedures: An extensional procedure strips away the representation and only considers the classical decision problem. An example of an extensional procedure is **maximizing** (Example 2.10). Note that **EXT** is a generalized version of *order invariance* as introduced in [10]. An intensional procedure, in contrast, considers only the representation. Hence, a choice is independent of the concrete alternatives and dependent only on the representation. Intensional procedures are, for instance, **First option on list** (Example 2.15), and **Second option on list** (Example 2.16). Examples 2.13, 2.12, 2.17, 2.18, and 2.19 are neither fully extensional nor fully intensional.

3.4 Relation between procedural properties

In the next section, we will use the classical properties, $\alpha\mathbf{E}$ and $\gamma^+\mathbf{E}$, to fully characterize procedures which can be rationalized. We will also consider the more procedural conditions, **SIND**, **TIIA**, **EXT**, and **INT**, and investigate when they are sufficient for a procedure to be rationalizable. To that end, we next investigate the relationship between $\alpha\mathbf{E}$ and $\gamma^+\mathbf{E}$ on the one side and **SIND**, **TIIA**, **EXT**, and **INT** on the other side. We begin with $\alpha\mathbf{E}$.

Proposition 3.4 Here are some facts about $\alpha\mathbf{E}$

(i) **INT** \Rightarrow $\alpha\mathbf{E}$.

(ii) **(EXT and SIND)** \Rightarrow $\alpha\mathbf{E}$.

(iii) For substitutable TX , **(SIND and TIIA)** \Rightarrow $\alpha\mathbf{E}$.

(iv) $\alpha\mathbf{E} \not\Rightarrow$ **SIND** and $\alpha\mathbf{E} \not\Rightarrow$ **INT**.

(v) For substitutable TX , **TIIA** $\not\Rightarrow$ $\alpha\mathbf{E}$.

(vi) For substitutable TX , $\alpha\mathbf{E} \not\Rightarrow$ **TIIA**.

Proof. (i) Let $x = P(a)$ and $x \in B \subseteq \llbracket a \rrbracket$. We need to show that there is a $b \in TX(\llbracket b \rrbracket = B \wedge x = P(b))$. Let b' be any structure such that $\llbracket b' \rrbracket = B$. Assume $y = P(b')$. Let $b = b'[x/y]$. By **INT** we have that $P(b) = P(b'[x/y]) = x$, since $P(b') = y$.

(ii) Given problem a on which P chooses x , re-represent a as a' such that B is a sub-problem, and is the only sub-problem containing x . Because P is **EXT**, it should still choose x . But since P is **SIND**, it has to choose x on that sub-problem, which gives a representation of B in which x is also chosen.

(iii) Assume TX is a substitutable ADT and suppose P satisfies **TIIA** and **SIND**. Given a problem $a \in TX$ such that $P(a) = x$, let us show how to remove an element from $\llbracket a \rrbracket$ (different from x) in such a way that x is still chosen. By induction we can then remove all the elements in $\llbracket a \rrbracket \setminus B$ obtaining $\alpha\mathbf{E}$. Let $y \in \llbracket a \rrbracket$ such that $x \neq y$ be the element we would like to remove. Since P satisfies **SIND**, we look at the bottom-up flow of choices, and identify the top-most point where y is not chosen (but some other element z is chosen). Consider the largest sub-problem a' such that $P(a') = y$. By definition we also have a larger problem $a'' = C \dots a' \dots$ in which z is chosen. Hence, by **TIIA** we can represent the set $(\llbracket a' \rrbracket \setminus \{y\}) \cup \{z\}$ as a problem b such that $P(C \dots b \dots) = z$. Finally, since P is **SIND**, replacing the sub-problem a'' by the new sub-problem $C \dots b \dots$ does not affect the overall choice x since $P(a'') = z = P(C \dots b \dots)$.

(iv) Example 2.16 (**Second option on list**) describes a procedure which is $\alpha\mathbf{E}$ but not **SIND**. Examples 2.10 and 2.11 describe procedures which satisfy $\alpha\mathbf{E}$ but not **INT**.

(v) As discussed in the proof of Proposition 3.3 (ii), Example 2.17 (**Larger sub-problem bias**) describes a procedure which is **TIIA**. It is easy to see, however, that this procedure is not $\alpha\mathbf{E}$. When $|\llbracket a \rrbracket| \geq N$ the procedure involves a mixture of intensional and extensional aspects, but in a sub-problem with $|\llbracket b \rrbracket| < N$ the procedure behaves extensionally.

(vi) Example 2.18 (**Smaller sub-problem bias**) describes a procedure which is $\alpha\mathbf{E}$ but not **TIIA**. It is clear that $\alpha\mathbf{E}$ holds, since making the problem smaller will essentially not change the behaviour of the procedure. To see that **TIIA** fails we consider the tree

$$a = \text{Node}(\text{Node } x (\text{Node } y z)) \underbrace{(\text{Node } v (\text{Node } w (\text{Node } u x)))}_{a'}$$

and suppose $N = 1$. In this case $P(a) = x$ and $P(a') = v$. But as soon as v replaces x in a' , there is no way to re-represent that sub-problem in such a way that x is still chosen overall. \square

Turning to $\gamma^+\mathbf{E}$, we have the following results.

Proposition 3.5 Here are some facts about $\gamma^+\mathbf{E}$

(i) **INT** \Rightarrow $\gamma^+\mathbf{E}$ but $\gamma^+\mathbf{E} \not\Rightarrow$ **INT**

(ii) If TX is substitutable then **SIND** \Rightarrow $\gamma^+\mathbf{E}$. But in general $\gamma^+\mathbf{E} \not\Rightarrow$ **SIND**.

(iii) **EXT** \cap $\gamma^+\mathbf{E} \neq \emptyset$, but **EXT** $\not\Rightarrow$ $\gamma^+\mathbf{E}$ and $\gamma^+\mathbf{E} \not\Rightarrow$ **EXT**

Proof. (i) If $P: TX \rightarrow X$ is intensional we can always place $x \in X$ in the position of the structure a that is chosen, so that $x = P(a)$. For a procedure which satisfies $\gamma^+\mathbf{E}$ but is not intensional we can simply take Example 2.10 (**Maximizing**).

(ii) Let $P: TX \rightarrow X$ be a strongly inductive procedure, i.e.

$$\forall C_i \forall x_1, \dots, x_{n_i} \in X \exists k \in \{1, \dots, n_i\} \forall a_1, \dots, a_{n_i} \left(\bigwedge_{j=1}^{n_i} x_j = a_j^* \rightarrow x_k = P(C_i a_1 \dots a_{n_i}) \right)$$

where $a_j^* = P(a_j)$ if $a_j \in TX$ and $a_j^* = a_j$ if $a_j \in X$. Let $a, b: TX$ be such that $x = P(a)$ and $y = P(b)$ and $y \in \llbracket a \rrbracket$. We must show that for some $d: TX$ with $\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket$ we have $x = P(d)$. Let $d = a[b/y]$. Clearly $\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket$. We prove that $x = P(a[b/y])$ by induction on the structure of the context $a[\cdot]$. If that is a constructor then

$$a[b/y] = C_i b_1 \dots b \dots b_{n_i}$$

so that

$$P(a[b/y]) = P(C_i b_1 \dots b \dots b_{n_i}) \in \{P(b_1), \dots, y, \dots, P(b_{n_i})\}.$$

But since $a = a[y/y]$ and $x = P(a)$ we must have $x = P(a[b/y])$. If $a[\cdot]$ is not a constructor then b appears inside a sub-context $b_j[\cdot]$, i.e.

$$a[b/y] = C_i b_1 \dots b_j[b/y] \dots b_{n_i}$$

but by induction hypothesis $P(b_j[b/y]) = P(b_j)$, so that $P(a[b/y]) = P(a) = x$. For a procedure that satisfies $\gamma^+ \mathbf{E}$ but not **SIND** consider Example 2.18 (**Small problem bias**). This procedure fails to be strongly inductive because it behaves differently depending on the size of the problem given. But it is still $\gamma^+ \mathbf{E}$ since it will choose in the same way when faced with a sub-problem which contains the chosen element.

(iii) The procedure $P: TX \rightarrow X$ from Example 2.10, which chooses the maximal element of $\llbracket a \rrbracket$ according to a fixed ordering of X , is both extensional and satisfies $\gamma^+ \mathbf{E}$. The procedure from Example 2.14 (Size-sensitive default element), satisfies **EXT** but $\gamma^+ \mathbf{E}$ fails. Finally, any procedure which satisfies **INT** will also satisfy $\gamma^+ \mathbf{E}$, but will not satisfy **EXT**. \square

Lastly, if a procedure is extensional, the following equivalence holds.

Proposition 3.6 *Suppose a procedure $P: TX \rightarrow X$ is **EXT**. Then the following statements are equivalent.*

(i) P satisfies **SIND**

(ii) P satisfies $\alpha \mathbf{E}$

(iii) P satisfies $\gamma^+ \mathbf{E}$

Proof. We will prove: (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)

(i) \Rightarrow (ii). Proposition 3.4 (ii).

(ii) \Rightarrow (iii). Let $P(a) = x$ and $P(b) = y$ and $y \in \llbracket a \rrbracket$. We need to show that there is a $d \in TX$, with $\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket$ and $P(d) = x$. Suppose $P(d) = z$. There are two cases:

1) $z \in \llbracket a \rrbracket$. Then, by $\alpha \mathbf{E}$ and extensionality we must have $z = P(a)$ and hence $z = x$.

2) $z \in \llbracket b \rrbracket$. Then, by $\alpha \mathbf{E}$ and extensionality $z = P(b)$ and then $z = y$. Hence $z \in \llbracket a \rrbracket$ and by the case above $z = x$.

(iii) \Rightarrow (i). Let $P(a_j) = x_j$ for some $a_1, \dots, a_{n_i} \in TX$. We need to show that for $b = C_i a_1 \dots a_{n_i}$ we must have $P(b) = x_j$ for some j . Let $z = P(b)$ with $z \in \llbracket a_j \rrbracket$. We have that $x_j = P(a_j)$ and $z \in \llbracket a_j \rrbracket$ and $z = P(b)$. Hence, by $\gamma^+ \mathbf{E}$ we must have a c such that $\llbracket c \rrbracket = \llbracket a_j \rrbracket \cup \llbracket b \rrbracket$ and $x_j = P(c)$. Since P is extensional $P(d) = P(b)$ and hence $x_j = z$. \square

If a procedure is extensional, it ignores the representation. Then, classical properties are equivalent. What is more, the equivalence extends to procedures which are strongly inductive, i.e. which proceed by a divide-and-conquer strategy.

4 Decision Procedures and Rational Choice

Our main objective in this paper is to characterize under which conditions a choice procedure can be rationalized by a preference relation. We will first work out which conditions guarantee that a procedure implements a (single-valued) choice function which can be rationalized. We will then turn our attention to (multi-valued) choice correspondences.

4.1 Choice functions

Under which conditions can choices made according to a procedure on represented decision problems be rationalized by a choice function? We begin with defining the relationship between a procedure and a choice function:

Definition 4.1 *A decision procedure $P: TX \rightarrow X$ is a choice function implementation (CFI) if its input-output behavior coincides with the canonical decision procedure for some choice function $c: 2^X \rightarrow X$, i.e.*

$$\exists c: 2^X \rightarrow X \forall a \in TX, P(a) = c(\llbracket a \rrbracket)$$

As the next proposition shows, **CFI** is intimately linked to extensionality **EXT**.

Proposition 4.2 CFI \Leftrightarrow EXT

Proof. Assume $P: TX \rightarrow X$ implements the choice function $c: 2^X \rightarrow X$, and that $a_0 \sim a_1$, i.e. $\llbracket a_0 \rrbracket = \llbracket a_1 \rrbracket$. Then $P(a_0) = c(\llbracket a_0 \rrbracket) = c(\llbracket a_1 \rrbracket) = P(a_1)$. For the other direction, assume P is extensional, and let $r: 2^X \rightarrow TX$ be a T -representation, i.e. a function such that $\llbracket r(A) \rrbracket = A$. It follows that $a \sim r(\llbracket a \rrbracket)$, and, by the extensionality of P , we have $P(a) = P(r(\llbracket a \rrbracket)) = \{\{P\}\}_r(\llbracket a \rrbracket)$. So P implements the choice function $\{\{P\}\}_r$. \square

We are interested in choice functions which can be rationalized. Formally,

Definition 4.3 (Rationalizable choice function) *A choice function $c: 2^X \rightarrow X$ is rationalizable if there exists a strict preference relation (X, \succ) such that for all $A \subseteq X$*

$$c(A) = \{x \mid \forall y \in A (x \succ y)\}$$

The strictness of the preference relation is required to make sure that the set $\{x \mid \forall y \in A (x \succ y)\}$ is a singleton.

We now turn to the definition of rationalizability in the context of a choice procedure:

Definition 4.4 (Rationalizable by choice function) *A decision procedure $P: TX \rightarrow X$ is said to be rationalizable by choice function if it satisfies (CFI) for a choice function $c: 2^X \rightarrow X$ which is rationalizable.*

In words, P is rationalizable by a choice function, if P implements a choice function whose choices agree with the maximization of a strict preference relation. Using $\alpha\mathbf{E}$ and Proposition 4.2, we obtain our first result:

Proposition 4.5 *P is rationalizable by choice function iff P is extensional and fulfills $\alpha\mathbf{E}$.*

Proof. “Left-to-right:” First note that if P is rationalizable by choice function, it has to fulfill **CFI**. Given Proposition 4.2 it thus has to be extensional.

Second, assume $x = P(a)$ and $x \in B \subseteq \llbracket a \rrbracket$. We need to show that $\exists b \in TX (\llbracket b \rrbracket = B \wedge x = P(b))$. If $x = P(a)$, then $c(\llbracket a \rrbracket) = x$. By definition of choice function rationality and $c(\llbracket a \rrbracket)$ it follows that $x \succsim y, \forall y \in \llbracket a \rrbracket$. But then, as $B \subseteq \llbracket a \rrbracket$, it also holds that $x \succsim y', \forall y' \in B$. And thus it follows

$c(B) = P(b) = x$ for any $\llbracket b \rrbracket = B$.

“Right-to-Left.” Assume P is extensional and satisfies $\alpha\mathbf{E}$. Define a binary relation $x \succsim y$ as follows

$$x \succsim y \quad \equiv \quad \forall a \in TX(\{x, y\} = \llbracket a \rrbracket \rightarrow x = P(a))$$

Since P is assumed to be extensional we can think of P as a mapping on 2^X , and $x \succsim y$ iff $x = P(\{x, y\})$. So the relation $x \succsim y$ is clearly complete and asymmetric: either $x = P(\{x, y\})$ or $y = P(\{x, y\})$. To show the transitivity of the relation, assume $x \succsim y$ and $y \succsim z$, i.e. $x = P(\{x, y\})$ and $y = P(\{y, z\})$. We need to show that $x = P(\{x, z\})$. We note that $y = P(\{x, y, z\})$ leads to a contradiction, by $\alpha\mathbf{E}$, since $\{x, y\} \subset \{x, y, z\}$ and $x = P(\{x, y\})$. Similarly, $z = P(\{x, y, z\})$ leads to a contradiction since $\{y, z\} \subset \{x, y, z\}$ and $y = P(\{y, z\})$. Hence, we must have $x = P(\{x, y, z\})$, which by $\alpha\mathbf{E}$ implies $x = P(\{x, z\})$. Lastly, we need to show that $P(A)$ is the \succsim -maximal element of A . Suppose $P(A) = x$ but $x \neq y = \max_{\succsim}(A)$. By $\alpha\mathbf{E}$ we must have that $x = P(\{x, y\})$, which implies $x \succsim y$, contradicting the assumption that y was maximal in A . \square

Extensionality is a strong property. It requires that the problem representation is completely ignored, and choices are made solely based on the set of values being represented. Thus, rationalizability by choice function is restrictive. Any procedure which is either intensional or at least has an intensional dimension, cannot be rationalized by a choice function. That also applies to the bulk of examples we introduced before.

The role of $\alpha\mathbf{E}$ in the proposition above mirrors very much its role in the rationalizability of classical choice functions. Yet, there is also a procedural dimension to the proposition above. In proposition 3.6 we showed that, under extensionality, classical consistency conditions as well as procedural consistency conditions are equivalent. Hence, we have the following result:

Corollary 4.6 *A procedure P which is both **EXT** and **SIND** is rationalizable by a choice function.*

Proof. Follows immediately from Proposition 3.6 and Proposition 4.5 \square

So, we can also check rationalizability from a procedural point of view. If an agent computes his choices by a divide-and-conquer strategy, his procedure is a candidate for being rationalizable.

We can apply this result to the satisficing procedure introduced in Example 2.11 (**Satisficing D**). It is easy to see that it fulfills **SIND**.⁵ So this satisficing procedure is a potential candidate for rationalizability. However, it does not fulfill **EXT**. To see this, consider two problems $[x, y, z]$ and $[x, z, y]$ and let $u(y), u(z) > u^* > u(x)$. Then, $P([x, y, z]) = y$ and $P([x, z, y]) = z$ while it holds $\llbracket [x, y, z] \rrbracket = \llbracket [x, z, y] \rrbracket$. Thus P is not extensional and henceforth in this general form it is not rationalizable by choice function.⁶

4.2 Choice correspondences

Choice function implementation is not the only way in which choices resulting from a procedure can be associated with preference relations. Recall, the exact, concrete representation of a decision problem is not known, only its type. Say, we observe the final choices made in an experiment. We know that choice problems were presented sequentially but we do not know in which order concretely. We can consider a *choice correspondence* which is induced by a choice procedure. The correspondence describes choices being made by the procedure on different representations of the same decision problem. For illustration, consider an example on lists. Suppose some procedure P selects $P([x, y, z]) = P([x, z, y]) = P([y, z, x]) = P([z, y, x]) = x$ and $P([y, x, z]) = P([z, x, y]) = y$. That is, when x is the first or last element it chooses x , otherwise it chooses y . All these lists are representations of the same set $\{x, y, z\}$. P induces a choice correspondence which chooses on the choice problem $A = \{x, y, z\}$ the elements $\{x, y\}$. This perspective on choice correspondences for decision procedures on lists was introduced in [10] (see also [1, 12] for similar

⁵It is also easy to check that it does fulfill $\alpha\mathbf{E}$: Consider any $a \in \mathbf{List} X$ with $x = P(a)$ and $x \in B \subseteq \llbracket a \rrbracket$. When we list the elements of B in a list $b \in \mathbf{List} X$ using the same order as in a , then $\llbracket b \rrbracket = B$ and $x = P(b)$. Hence, $\alpha\mathbf{E}$ holds.

⁶In Section 5.1 where we consider extensions of our framework, we show that one can impose more structure on the representation such that a rationalization by choice function is possible.

interpretations). Here, we consider such an interpretation not only for lists but for arbitrary representations by algebraic datatypes. We will investigate under which conditions the induced correspondence can be rationalized.

We begin with defining a choice correspondence induced by a procedure:

Definition 4.7 (Choice correspondence of a decision procedure) *Given a decision procedure $P: TX \rightarrow X$, we define its associated choice correspondence $C_P: 2^X \rightarrow 2^X$ as*

$$C_P(A) = \{P(a) \mid a \in TX \text{ such that } \llbracket a \rrbracket = A\}$$

In words, given a decision procedure and a standard decision problem $A \subseteq X$, we consider all possible ways of representing the set A as a T -structure, and observe what the decision procedure would choose in each case. This gives us a set of choices which we use to form the choice correspondence.

In the example above, the procedure $P: \mathbf{List} X \rightarrow X$ gives rise to the correspondence C_P which on the set $\{x, y, z\}$ chooses $C_P(\{x, y, z\}) = \{x, y\}$. Thus, the choice correspondence C_P can be seen as an approximation of the actual decision-making procedure P .

Next, we formally introduce rationalizability by choice correspondence:

Definition 4.8 (Choice correspondence rationalizable) *A choice correspondence $C: 2^X \rightarrow 2^X$ is rationalizable, if there exists a preference relation (X, \succeq) , that is, a complete and transitive binary relation \succeq on X , such that*

$$C(A) = \{x \mid \forall y \in A (x \succeq y)\}$$

A procedure $P: TX \rightarrow X$ is choice correspondence rationalizable if C_P is rationalizable.

Our aim is to characterize procedures which are choice correspondence rationalizable.

Proposition 4.9 *For any algebraic datatype TX and procedure $P: TX \rightarrow X$, the choice correspondence C_P is rationalizable iff P satisfies $\gamma^+ \mathbf{E}$ and $\alpha \mathbf{E}$.*

Proof. Assume first that C_P is rationalizable, and that $x \in P(a)$ and $y \in P(b)$ and $y \in \llbracket a \rrbracket$. For $\gamma^+ \mathbf{E}$ we need to show

$$\exists d \in TX (\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket \wedge x = P(d))$$

Our assumptions imply that $x \in C_P(\llbracket a \rrbracket)$ and $y \in C_P(\llbracket b \rrbracket)$. Since C_P is rationalizable, we know that x is one of the maximal elements in $\llbracket a \rrbracket$ and y a maximal element in $\llbracket b \rrbracket$. Moreover, since $y \in \llbracket a \rrbracket$, we have that $x \succeq y$, so that, by transitivity, x is also maximal in $\llbracket a \rrbracket \cup \llbracket b \rrbracket$, i.e. $x \in C_P(\llbracket a \rrbracket \cup \llbracket b \rrbracket)$. Finally, by the definition of C_P there exists a representation d such that $\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket$ with $x = P(d)$. This shows that P satisfies $\gamma^+ \mathbf{E}$. In order to prove $\alpha \mathbf{E}$, assume again that C_P is rationalizable, $x = P(a)$, and $x \in B \subseteq \llbracket a \rrbracket$. We must show that for some b we have $\llbracket b \rrbracket = B$ and $x = P(b)$. But that is clear since if x is a maximal element in $\llbracket a \rrbracket$ it must also be a maximal element in b for any b such that $\llbracket b \rrbracket \subseteq \llbracket a \rrbracket$.

For the other direction, assume P satisfies both $\gamma^+ \mathbf{E}$ and $\alpha \mathbf{E}$. Define a binary relation $x \succeq y$ as follows

$$x \succeq y \quad \equiv \quad \exists a \in TX (x, y \in \llbracket a \rrbracket \wedge x = P(a))$$

Clearly this relation is total. In order to show it is a preference relation, it remains to verify it is transitive. Hence, assume $x \succeq y$ and $y \succeq z$, i.e.

$$\exists a \in TX (x, y \in \llbracket a \rrbracket \wedge x = P(a)) \quad \text{and} \quad \exists b \in TX (y, z \in \llbracket b \rrbracket \wedge y = P(b))$$

By $\gamma^+ \mathbf{E}$ we have that $\exists d \in TX (\llbracket d \rrbracket = \llbracket a \rrbracket \cup \llbracket b \rrbracket \wedge x = P(d))$. Since $z \in \llbracket d \rrbracket$, we are done. It remains to show that $x \in C_P(A)$ if and only if $x \in \max_{\succeq}(A)$. If $x \in C_P(A)$ then $x = P(a)$ for some $\llbracket a \rrbracket = A$. And by definition of \succeq we do have that x is a maximal element in A , i.e. $x \succeq y$ for all $y \in A$. For the other direction, assume $x \in \max_{\succeq}(A)$. Let $A = \{x, y_1, \dots, y_n\}$. Since $x \succeq y_i$ for all $1 \leq i \leq n$, we have a sequence a_i such that $x = P(a_i)$ and $\{x, y_i\} \subseteq \llbracket a_i \rrbracket$. By $\gamma^+ \mathbf{E}$ there is a d such that $\llbracket d \rrbracket = \bigcup_{i=1}^n \llbracket a_i \rrbracket$ and $x = P(d)$. Finally, by $\alpha \mathbf{E}$ there is a d' such that $\llbracket d' \rrbracket = \{x, y_1, \dots, y_n\}$ such that $x = P(d')$. \square

This proposition fully characterizes when the correspondence induced by a procedure P can be rationalized. It is easy to see that the Satisficing example (Example 2.11) fulfills $\gamma^+\mathbf{E}$. As shown before, because it is not extensional, Example 2.11 fails to be rationalized by choice *function*. Yet, as it satisfies $\alpha\mathbf{E}$, by the last proposition it is rationalizable by choice *correspondence*.

Beyond the characterizations above, what do we learn by introducing our machinery? In the section before, we already showed (Proposition 3.4) that procedures which satisfy \mathbf{INT} also satisfy $\alpha\mathbf{E}$. Moreover, by Proposition 3.5, \mathbf{INT} also implies $\gamma^+\mathbf{E}$. The following result is therefore immediate:

Corollary 4.10 *Procedures $P: TX \rightarrow X$ which satisfy \mathbf{INT} are choice correspondence rationalizable.*

Proof. Follows immediately from Propositions 3.4, 3.5, and 4.9 □

Hence, an intensional procedure, which exclusively takes the representation of the problem into account and does not care about the actual values of elements, is rationalized by choice correspondence. This may seem surprising at first. Yet, the reason is simple: For procedures $P: TX \rightarrow X$ satisfying \mathbf{INT} we have that $C_P: 2^X \rightarrow 2^X$ is the identity function, since we can always change the representation to force an element to be chosen. Hence, C_P can be rationalized by the preference relation which rates all elements on equal footing.

Intensionality is not the only procedural property that is related to choice correspondence rationalization. Our discussion in Section 3.4 showed that procedural consistency is sufficient to fulfill classical consistency. We therefore have another immediate corollary:

Corollary 4.11 *For a substitutable ADT, a procedure P which is \mathbf{SIND} and T -independent of irrelevant alternatives is rationalized by choice correspondence.*

Proof. Follows immediately from Propositions 3.4, 3.5, and 4.9 □

5 Representations and Meaning

Intensional procedures beg an obvious question: How do we interpret a preference relation that was derived from an intensional procedure? If we just observe the set of alternatives and the choices, we may accept that the agent chose according to a preference relation even though his procedure was independent of the concrete alternatives. It appears that such preferences do not carry any economic meaning.

But this raises another immediate question: Does such a procedure make sense at all and why should an individual choose in this way? Our answer is: Yes, such procedures make sense, *if* we consider additional information that often is available for an agent. We consider such an extension in this section. So far, we have considered algebraic datatypes as syntactical constructs that an agent's procedure can access when making a decision. In the following we consider an extension of this basic framework – representations that convey economic meaning. Up until now X has been simply taken to be a *set* of alternatives. We now look at the cases when X is itself the Cartesian product of other sets $X = X_1 \times \dots \times X_n$ (thinking of each component as an attribute of X), or a set of sets $X = \mathcal{P}(Y)$ (thinking of agents *filtering* a sub-set of alternatives) or both $X = \mathcal{P}(X_1 \times \dots \times X_n)$ (filtering based on attributes). Lastly, we consider the case where X contains representations, i.e. another layer of ADTs. At first, the last aspect may appear as a technicality, whereas in fact it is of practical relevance for intuitive ways of decision-making.

5.1 Representations conveying problem structure

Consider first the case when X is assumed to carry some information itself. Often people care about the properties goods have. For instance, when choosing a wine in a restaurant, an agent may not just consider a simple set of wines W , but rather certain *attributes* of wines such as grape G , year Y , price P and country of origin C . So the set of wines W is rather a set of tuples $W = G \times Y \times P \times C$ characterizing the different

wines available. When we represent wines as an ADT, say **List** W we might want to group all wines of the same country together, or we might want to list the wines in order of price. These can be seen as *guarantees* on the representation of the problem. Take the example of wines listed by price. If the decision-maker has information about this guarantee, he can assume that only lists l : **List** W which satisfy

$$i < j \Rightarrow (l_i)_P < (l_j)_P$$

will be considered, where x_P denotes the P coordinate of x : $G \times Y \times P \times C$.

This guarantee on the representation of the problem is important because an agent may make use of this information in his decision procedure. Say an agent really likes white burgundies from a certain region and he is presented a list of recommendations which contains wines ordered according to these criteria. He may then simply choose the first element of the list. Contrast this with the case where he faces a list of wines that are not ordered according to his criteria. Then he might need to scan the whole list in order to make a choice. Another possible guarantee could be that a representation is guaranteed to contain no duplicates. This could also affect how an agent chooses. Consider a procedure where the agent decides according to the number of appearances of an element (he chooses the element which appears most often). If there is no duplication, this would of course change his behavior and in this case simplify his procedure (to always choosing the first element). Overall, guarantees are crucial because they allow different ways of how an agent can choose and what the economic interpretation of such procedures is.

What guarantees make sense? This depends on the situation to be modeled and cannot be answered in general. Because of that, in the sections before for a given ADT, we considered *all* representations possible for this set. This is the minimum we require. Any guarantee imposed beyond that – which may make sense from a modeler’s perspective in a particular situation – will restrict this set of representations.

We discussed before that intensional procedures can be rationalized but that the meaning of preferences derived from such procedures is not clear. Once we attach guarantees to representations, intensional procedures are less “irrational”. Choosing the first alternative, for instance, may make sense and could be a particularly simple and effective way of choosing.⁷ We will consider an example below.

Besides affecting the agent’s procedure, guarantees also affect the modeler’s ability to rationalize choices. Recall the **Satisficing** procedure introduced in Example 2.11. It is not rationalizable by choice function because it is not extensional. Consider now the following adapted version which comes with a guarantee.

Example 5.1 (Satisficing II) *Consider a variant of the satisficing **List**-decision procedure given in Example 2.11, whereby the alternatives are always listed according to some predetermined order $<$ on X .*

This guarantee, here a restriction on the domain of lists, ensures that this satisficing procedure is extensional on the restricted domain of lists that satisfy the global ordering $(X, <)$. Indeed, on such lists we have that P is extensional as it holds

$$a_0 \sim a_1 \Rightarrow Pa_0 = Pa_1$$

As P does fulfill $\alpha\mathbf{E}$, this satisficing procedure on the restricted domain can be rationalized by choice function. The fact that the above version of satisficing can be rationalized if more constrains are put on the sequence of alternatives is discussed in [9, p.46]. The above example shows that one can interpret his result as a special form of guarantee.

5.2 Choosing by checklist

Continuing with the example of an agent choosing a wine, suppose agents are given two main criteria to narrow down their search: whether it is a red wine, $R = \mathbb{B}$, and whether it is dry, $D = \mathbb{B}$, where \mathbb{B} denotes the two-element set of Booleans. In this simplistic model the set of wines is given by $W = G \times Y \times R \times D$, with G denoting the set of grapes, and Y the set of years. Using these two criteria agents can filter a set of

⁷A likely driving force behind such issues, which we do not pursue in the current paper, are complexity issues. Just choosing the first element of a list is very simple, maximizing according to some order is not. Thus an external agent which can represent alternatives in a certain way may simplify substantially the choice process of an agent.

wines satisfying their preferences. The four sub-sets of wines are best represented as a labeled binary tree $\mathbf{LTree} X$ with two labels

$$\text{Wines } X \Rightarrow \text{Wine } X \mid \text{Node}_{\text{red?}} (\text{Wines } X) (\text{Wines } X) \mid \text{Node}_{\text{dry?}} (\text{Wines } X) (\text{Wines } X)$$

If W is the set of all wines, and $\mathcal{P}(W)$ the power-set of the set W , then the represented decision problem would be $\mathbf{LTree}(\mathcal{P}(W))$ – navigating the tree will lead the agent to a set of wines matching the chosen criteria.

Assuming an agent is aware of the meaning attached to a representation (as it is the case with online shops which allow customers to filter search results), an agent who wishes to filter non-dry red wines will be able to easily navigate the represented decision problem as

$$P(\text{Wines } W) = W$$

$$P(\text{Node}_{\text{red?}} a b) = P(a)$$

$$P(\text{Node}_{\text{dry?}} a b) = P(b)$$

In [4] it is shown that the choice correspondence deriving from “choosing by checklist” can be rationalized. As the example above shows, choosing by checklist can be represented as a decision procedure on finitely branching binary trees which is intensional, since the choice of the final elements solely depends on the position within the tree. The fact that such a procedure can be rationalized follows directly from Corollary 4.10.

5.3 Nested ADTs

In the choosing by checklist procedure above, an agent is navigating a tree and selecting a *set* of elements that contain only elements with certain properties. We can go one step further and consider the representation of these subsets as e.g. *lists*, and let the agent choose one of these elements from the list. The overall procedure has type

$$\mathbf{LTree} (\mathbf{List} W) \rightarrow W$$

i.e. the decision maker is choosing a single wine when given a tree of lists of wines. Considering that wines also have a *price* property, and assuming that each of the sub-lists are guaranteed to be given by an increasing price order, we can combine procedures we have seen before:

$$P(\text{Wines } L) = \text{first}(L)$$

$$P(\text{Node}_{\text{red?}} a b) = P(a)$$

$$P(\text{Node}_{\text{dry?}} a b) = P(b)$$

is the agent who chooses the cheapest non-dry red wine. This nested structure of trees of lists is indeed common in online shopping. For instance, as the consumers apply different filters on the Amazon website, they are presented with a *list* of sub-results that are guaranteed to match their filter selection. The more filters one applies the shorter the list of results one gets. Moreover, these lists are normally ordered by some criteria such as price (increasing or decreasing), ratings or relevance.

We began our paper with a specific representation of search results. We can think of this as another nested type. Let us consider this here. We begin with the most primitive one, a single item

$$\mathbf{Item} X \Rightarrow \mid X$$

Next, we consider a page, which contains ten items:

$$\mathbf{Page} X \Rightarrow \text{P } \mathbf{Item}(X)_1 \dots \mathbf{Item}(X)_{10}$$

and lastly, the search result is constructed as a first page and further results:

$$\mathbf{Result\ X} \Rightarrow R_1 \mathbf{Page(X)} \mid R_2 \mathbf{Page(X)} \mathbf{Result(X)}$$

Search results are interesting because they often come with further guarantees. Another interesting example of a nested ADT that comes with guarantees is the results of a search engine (e.g. Google or Bing). The results of the search are represented as **Result**. But the overall items are organized in decreasing order of page rank. The actual formula for calculating the rank of a page is a trade secret. Nevertheless, people will often choose the first link of the first page, because they have learned to trust that the ranking indeed gives them the best match to their search keywords. In modeling terms we have again a nested ADT. A person making a search and choosing one of the given links will be using a procedure of type

$$\mathbf{Result\ X} \rightarrow \mathbf{X}$$

The interesting phenomena is that what might look like a completely *intensional* procedure, of choosing the first link on the first page, is in fact *extensional* once we have the guarantee of the ordering in which the search results are given: The person is indeed choosing the page with the highest rank!

The simple extensions we have introduced in this section should illustrate that the language of ADT's and procedures is expressive enough to capture various aspects of representations. Moreover, the theory introduced before for the more limited setting seamlessly extends to this enriched environment. What these extensions should also make clear is that the economic meaning of a given procedure crucially depends on the context in which it is embedded. Choosing the first element of a list at first sight may strike one as a strange idea. Yet, if not seen in isolation but as part of a richer environment where the list carries meaning or as part of a nested procedure, it becomes reasonable.

6 Conclusion

In this paper we model the constraints agents underlie when choosing. Alternatives often can only be accessed in a predetermined way. Online webshops, for instance, design the way agents can search for goods. We use an algebraic structure, a context-free grammar, to model the restrictions choice procedures have to respect. We characterize procedures that can be rationalized by choice functions and choice correspondences induced by the represented decision problem. Moreover, we show that there are procedural consistency conditions which are sufficient for a rationalization to exist. While intensional procedures may look irrational at first, it is important to view them in context. We provided an example in the last section which shows that even for a perfectly rational economic agent circumstances exist where they make perfect sense.

The language we introduce opens several immediate avenues for future questions. In the literature on bounded rationality, some researchers have considered decision makers as automata to restrict the complexity of their decision-making process, see for instance [11]. In this paper we considered decision procedures defined by primitive recursive functions. We could consider other complexity classes of functions. Moreover, the fact that representations may help agents to make decisions effectively is important. As illustrated in the last section, in the extreme case they can exclusively rely on the representation. When assessing how hard it is for an agent to decide, this has to be taken into account. To that end it is in particular important to consider composite procedures - which is possible within the framework we have introduced.

We defined a represented decision problem as the primitive and we assumed that agents operate on this representation. However, an agent might abstract a represented decision problem by ignoring some aspects of the structure, e.g. given a list the agent might not care about the order but might look at how often each element appears. An agent may also transform a represented decision problem into another representation and apply a procedure on that representation.

References

- [1] Douglas B. Bernheim and Antonio Rangel. Beyond Revealed Preference : Choice-Theoretic Foundations for Behavioral Welfare Economics. *The Quarterly Journal of Economics*, 124(1):51–104, 2009.

- [2] Andrew Caplin and Mark Dean. Search, choice, and revealed preference. *Theoretical Economics*, 6(1):19–48, 2011.
- [3] Dinko Dimitrov, Saptarshi Mukherjee, and Nozomu Muto. divide-and-choose in list-based decision problems. *Theory and Decision*, 81(1):17–31, 2016.
- [4] Michael Mandler, Paola Manzini, and Marco Mariotti. A million answers to twenty questions : Choosing by checklist. *Journal of Economic Theory*, 147(1):71–92, 2012.
- [5] Yusufcan Masatlioglu and Daisuke Nakajima. Choice by iterative search. *Theoretical Economics*, 8(3):701–728, 2013.
- [6] Saptarshi Mukherjee. Choice in ordered-tree-based decision problems. *Social Choice and Welfare*, 43(2):471–496, 2014.
- [7] Bryan O’Sullivan, Don Stewart, and John Goerzen. *Real World Haskell*. 2008.
- [8] Charles R. Plott. Path independence, rationality, and social choice. *Econometrica*, 41(6):1075–1091, 1973.
- [9] Ariel Rubinstein. *Lecture Notes in Microeconomic Theory*. SUNY-Oswego, Department of Economics, 2016.
- [10] Ariel Rubinstein and Yuval Salant. A model of choice from lists. *Theoretical Economics*, 1:3–17, 2006.
- [11] Yuval Salant. Procedural Analysis of Choice Rules with Applications to Bounded Rationality. *American Economic Review*, 101:724–748, 2011.
- [12] Yuval Salant and Ariel Rubinstein. (A,f) : Choice with Frames. *Review of Economic Studies*, 75:1287–1296, 2008.
- [13] Amartya K Sen. Choice functions and revealed preference. *The Review of Economic Studies*, 38(3):307–317, 1971.
- [14] Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955.
- [15] Herbert A Simon. Rational choice and the structure of the environment. *Psychological review*, 63(2):129, 1956.
- [16] Kemal Yildiz. List-rationalizable choice. *Theoretical Economics*, 11(2):587–599, 2016.

7 Appendix

In the following we provide the Haskell code for the ADTs as well as the choice procedures we use in this paper.

7.1 ADTs

```
-- List
data List x = Sing x | Cons x (List x)

-- List2
data List2 x = Sing2 x | Cat (List2 x) (List2 x)

-- Tree
```



```

data Tree x = Leaf x | Node (Tree x) (Tree x) (Tree x)

-- LTree
data LTree x = LLeaf x
             | Node1 (LTree x) (LTree x)
             | Node2 (LTree x) (LTree x)
             | Node3 (LTree x) (LTree x)

-- Wines
data Wines x = Wine x | Red (Wines x) (Wines x) | Dry (Wines x) (Wines x)

-- Web search

data Item x = I x

data Page x = P (Item x) (Item x) (Item x) (Item x)
             (Item x) (Item x) (Item x) (Item x) (Item x) (Item x)

data Result x = R1 (Page x) | R2 (Page x) (Result x)

```

7.2 Choice procedures

```

-- First element on List
first :: List x -> x
first (Sing x) = x
first (Cons x xs) = x

-- Leftmost element on Tree
left_most :: Tree x -> x
left_most (Leaf x) = x
left_most (Node t1 t2 t3) = left_most t1

-- Maximizing on list
best :: (Ord x) => List x -> x
best (Sing x) = x
best (Cons x xs) = let y = best xs in (if x > y then x else y)

-- Satisficing
sat :: (x -> Float) -> Float -> List x -> x
sat u t (Sing x) = x
sat u t (Cons x xs) | u x >= t = x
                   | otherwise = sat u t xs

--List2 Satisficing
sat2 :: (x -> Float) -> Float -> List2 x -> x
sat2 u t (Sing2 x) = x
sat2 u t (Cat t1 tr) | u left >= t = left

```

```

        | otherwise = sat2 u t tr
where left = sat2 u t tl

-- Conditional List satisficing

sat3 :: Ord x => [x] -> List x -> x
sat3 known xs | 2 * size_known > size_all = sat (value known) 0.5 xs
              | otherwise = best xs
  where
    size_all = count (\x -> True) xs
    size_known = count (\x -> x `elem` known) xs

-- Choose default on large problems
default2 :: Ord x => Int -> x -> List x -> x
default2 n d xs | d `elem` xs && xs_size > n = d
                | otherwise = best xs
  where xs_size = count (\x -> True) xs

-- First option on list2

first_2 :: List2 x -> x
first_2 (Sing2 v) = v
first_2 (Cat v1 vr) = first_2 v1

-- Second option on list

second :: List x -> x
second (Sing x) = x
second (Cons x (Cons y ys)) = y

-- Second option on list2

second_2 :: List2 x -> x
second_2 (Sing2 v) = v
second_2 (Cat (Sing2 v) vr) = second_2 vr
second_2 (Cat v1 vr) = second_2 v1

-- Larger sub-problem bias

bias_large :: Ord x => Int -> Tree x -> x
bias_large _ (Leaf v) = v
bias_large n (Node t1 t2 t3) | size' < n = best' (Node t1 t2 t3)
                             | s1 > max s2 s3 = bias_large n t1
                             | s2 > max s1 s3 = bias_large n t2
                             | otherwise = bias_large n t3
  where { size' = size (Node t1 t2 t3)
        ; s1 = size t1; s2 = size t2; s3 = size t3 }

```

```

-- Smaller sub-problem bias
bias_small _ (Leaf v) = v
bias_small n (Node t1 t2 t3) | size' < n = best' (Node t1 t2 t3)
                             | s1 < min s2 s3 = bias_small n t1
                             | s2 < min s1 s3 = bias_small n t2
                             | otherwise      = bias_small n t3
  where { size' = size (Node t1 t2 t3)
        ; s1 = size t1; s2 = size t2; s3 = size t3 }

-- Avoiding undesirable element
avoid :: Eq x => x -> Int -> List2 x -> x
avoid u n (Sing2 v) = v
avoid u n (Cat al ar) | avoid u n al /= u || size' >= n = avoid u n al
                     | otherwise = avoid u n ar
  where size' = sizeList2 (Cat al ar)

-- Checklist on wines
agent_wine :: Wines x -> x
agent_wine (Wine x) = x
agent_wine (Red x y) = agent_wine y
agent_wine (Dry x y) = agent_wine x

-- Checklist on list of wines
agent_wine2 :: Wines (List x) -> x
agent_wine2 (Wine xs) = first_1 x
agent_wine2 (Red x y) = agent_wine2 y
agent_wine2 (Dry x y) = agent_wine2 x

-- An example of a search result using x = Int

p1 :: Page Int
p1 = P (I 1) (I 2) (I 3) (I 4) (I 5) (I 6) (I 7) (I 8) (I 9) (I 10)

p2 :: Page Int
p2 = P (I 11) (I 12) (I 13) (I 14) (I 15) (I 16) (I 17) (I 18) (I 19) (I 20)

r1 :: Result Int
r1 = R2 p1 (R1 p2)

-- A procedure that selects the first element of the page
first_of_page :: Page Int -> Int
first_of_page (P (I n) (I _) (I _) (I _) (I _) (I _) (I _) (I _) (I _) (I _)) = n

first_of_search :: Result Int -> Int
first_of_search (R1 p) = first_of_page p
first_of_search (R2 p ps) = first_of_page p

```