# Formalising Fairness in the Assignment Problem with Ordinal Preferences in Isabelle/HOL

Duc Minh Do[1] and Christine Rizkallah[1]

The University of Melbourne, Melbourne, Victoria, Australia
{do.100, christine.rizkallah}@unimelb.edu.au

**Abstract.** Social choice theory is a multidisciplinary research area that studies collective decision-making based on individual preferences. A central problem in this domain is the fair assignment of indivisible resources, which arises in applications such as organ matching, school admissions, and job placements. Notions of fairness, such as proportionality and envy-freeness, are well-established; they are sometimes unattainable, particularly in the discrete setting, where resources cannot be divided. We present a mechanised formalisation of key fairness concepts for the assignment problem under ordinal preferences using the Isabelle/HOL theorem prover. Our development captures both discrete and fractional assignment settings and formalises the responsive set extension, which is a central concept for lifting preferences over resources to preferences over sets of resources, enabling the comparison of allocations under ordinal preferences. We formalise multiple fairness notions—including stochastic dominance (SD) proportionality and envy-freeness, as well as their relaxed variants—and we formally verify relationships between these notions. The formalisation builds on existing verified results in social choice theory and closes gaps found in pen-and-paper proofs.

**Keywords:** Social choice theory · Fairness · Interactive theorem proving

## 1 Introduction

Social choice theory is a multidisciplinary field that studies how individual preferences can be aggregated into collective decisions. It draws on ideas from economics, mathematics, political science, computer science, and philosophy. Despite its broad scope, problems in social choice theory share the following central question: how can individual preferences be combined into decisions that best reflect society's interests as a whole? A well-studied subarea of social choice theory is *voting theory*. Voting theory examines how to aggregate individual preferences over candidates into a collective outcome, such as a single winning candidate, a set of winning candidates, or a ranked list of the candidates [33]. It addresses fundamental questions about different voting systems, exploring how different voting mechanisms can produce different societal outcomes from the same list of individual preferences. Foundational results in this area, such as Arrow's impossibility theorem [2], the Condorcet Paradox [21], and the Gibbard–Satterthwaite theorem [38], highlight the inherent limitations of designing voting mechanisms that meet a small set of intuitive and seemingly reasonable criteria.

Another key subarea–central to this work–is the *assignment problem*, which involves allocating resources to individuals based on their preferences. This problem is also known as the *house allocation problem* and as the *two-sided matching with one-sided preferences* [5]. Here, the main challenge is to allocate a limited set of resources to individuals in a way that respects various levels of fairness and efficiency criteria. As such, this problem is sometimes also referred to as the *fair allocation problem* [12].

There are two main settings in which this problem is studied, based on whether or not the resources are divisible. In the case when the resources are divisible, the problem can be framed as a *fractional assignment* problem, where each individual receives a proportion of each resource. In the case when resources are indivisible, solutions typically involve either *discrete assignments*—where individuals receive whole resources—or *randomised assignments*—where individuals are assigned probabilities over whether or not they are allocated each resource [1].

Fairness in the assignment problem has real-world applications such as CPU time allocation, school admissions, job placements, public housing, and allocating organs to patients. Equitable resource allocation is critical in such scenarios. This underscores the importance of rigorously defining and reasoning about fairness.

Two of the most fundamental and widely studied notions of fairness are *envy-freeness* [19], and *proportionality* (only defined in the fractional and the randomised setting) [29, 16]. Envy-freeness requires that no individual prefers another individual's allocation to their own. Proportionality imposes a minimal requirement on each individual's allocation, requiring that each individual's allocation be at least as good as the uniform distribution of resources to individuals [37, 9].

While desirable, these fairness criteria are sometimes impossible to achieve in practice, especially in the discrete setting with indivisible resources [1, 3]. As a result, researchers study weaker variants such as *weak envy-freeness*, *weak proportionality*, and *proportionality up to a fixed number* of resources [9, 15].

As many of the foundational results in social choice theory are impossibility theorems and due to the field's amenability to satisfying desired axiomatic properties, there is a strong motivation for researchers to use theorem provers as a tool to aid the research process [4]. There has been an extensive line of work in verifying existing social choice theoretic results in proof assistants [27, 32, 31, 35]. Beyond verifying existing results, theorem provers and SAT solvers have also enabled researchers to quickly discover new theorems and insights in the field [7, 13, 14, 8]. Recent efforts have also focused on formalising core concepts in social choice theory, establishing a reusable foundation for future mechanisation efforts in this field [36, 34, 18]. Although important concepts in randomised social choice, such as stochastic dominance and utility functions, have been formalised in Isabelle/HOL, to the best of our knowledge, there is no general formalisation of the assignment problem or its associated fairness criteria.

This paper describes an Isabelle/HOL formalisation of the assignment problem that relates various notions of fairness in the *fractional/randomised* and *discrete* settings, and provides formal proofs of known relationships between these notions [6]. In particular, we define envy-freeness and proportionality using the stochastic dominance (SD) relation and using utility functions to compare allocations. For the discrete case, we also formalise the established notion of a responsive set extension (RS) relation, which lifts ordinal preferences over resources to preferences over sets of resources [6].

Developing a formalisation of the assignment problem establishes a theoretical foundation that enables verifying relationships between existing notions of fairness. While we verified most of the pen-and-paper results [6] relating these notions, one relationship between two notions of fairness turned out to be incorrect. For this notion, we provide a formalised counter-example.

We start by presenting some background on the theoretical foundations of the assignment problem and its fairness definitions, followed by an explanation of their mechanised counterparts in Isabelle/HOL. The formalisation is publicly available [17].

## 2 Preliminaries

This section introduces some background that is necessary for understanding our development. We begin by defining the assignment problem in both its fractional and discrete settings, followed by a presentation of the key fairness notions that we formalise. Our definitions and proofs follow existing pen-and-paper results [6].

**The Assignment Problem.** The assignment problem is typically defined as a triple $(N, O, \mathcal{R})$, where $N = \{a_1, a_2, \ldots, a_n\}$ is a finite set of agents, $O = \{o_1, o_2, \ldots, o_m\}$ is a finite set of objects, and $\mathcal{R} = \{R_1, R_2, \ldots, R_n\}$ is a profile of ordinal preference relations—one for each agent. Each relation $R_i$ is a weak total preorder over $O$, representing the preferences of agent $i$. An assignment $p$ is typically represented as an $n \times m$ matrix, where entry $p[i, j]$ indicates the fraction (or probability) that object $o_j$ is allocated to agent $a_i$. In the randomised setting, these entries are real numbers in [0, 1]. A common assumption in this setting is that there is only one occurrence of each object, thus, the sum of the probabilities of receiving a particular object over all agents must sum to one: $\sum_{i=1}^{n} p[i, j] = 1$, for all $j \in \{1...m\}$. In the discrete setting, with indivisible objects, the definition of an allocation is similar to that of the randomised setting, but with the added restriction that entries have to be either 0 or 1, depending on whether an agent receives a given object in full. To accommodate scenarios where agents may receive multiple objects, we relax the usual assumption that each agent's allocation is a lottery (i.e., sums to 1). Instead, we allow each agent's total allocation to sum to a fixed constant, which is shared across agents. This enables the use of the stochastic dominance even in scenarios when allocations are not lotteries.

**Fairness Notions.** As discussed earlier, various existing notions of fairness are

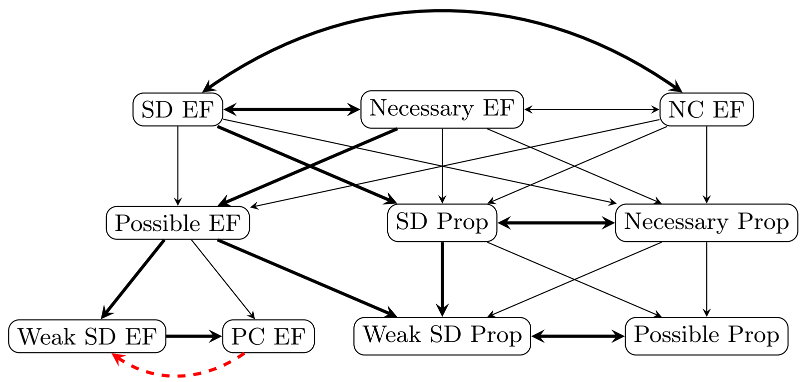


Fig. 1: Logical relationships between fairness notions as presented in prior work [6]. Envy-freeness is abbreviated EF, proportionality is abbreviated Prop, arrows indicate implication, and bidirectional arrows indicate equivalence. We use a dotted arrow to identify the implication we disproved and use bold to identify relationships that are verified directly; the remaining arrows are verified as corollaries.

studied in the literature. These notions form a hierarchy of fairness concepts [6], presented in Figure 1, that we formalise and verify in this paper.

*SD Based.* First, let's consider the two key notions of fairness based on the probabilistic *stochastic dominance* (SD) relation which compares preference allocations. An allocation satisfies *SD envy-freeness* (SD EF) if every agent prefers their allocation to that of any other agent under the SD relation. A weaker variant of this notion is *weak SD envy-freeness* (Weak SD EF), satisfied when no agent prefers other agents' allocation to their own. The second notion is *SD proportionality* (SD Prop), which is satisfied if every agent prefers their allocation to the uniform allocation of objects to agents. There is also a weaker variant of this notion, *weak SD proportionality* (Weak SD Prop), of which an assignment satisfies when no agent prefers the uniform allocation to their own.

*Utility-Based Variants.* We also formalise *utility-based* variants of each of these two fairness notions: *necessary fairness* holds if the fairness notion is satisfied for all utility functions that are consistent with each agent's ordinal preferences. Meanwhile, *possible fairness* holds if there exists at least one utility function for which the fairness notion is satisfied. The notion of *necessary fairness* includes both *necessary envy-freeness* (Necessary EF) and *necessary proportionality* (Necessary Prop), and the notion *possible fairness* includes both *possible envy-freeness* (Possible EF) and *possible proportionality* (Possible Prop).

*Responsive Set Extension Variants.* In the discrete setting, where allocations are sets rather than distributions, we lift preferences over individual objects to preferences over sets using the *responsive set extension* [10]. Given a preference relation $R$ over objects, an agent prefers a set $A$ to a set $B$ if there exists an injection $f : B \rightarrow A$ such that for all $x \in B$, $f(x)$ is preferred to $x$ under $R$. This provides a principled way to compare their allocated set of indivisible objects using only ordinal information. This lifting is used to define two additional fairness notions that have been studied in the discrete setting [10]. In particular, *necessary completion envy-freeness* (NC EF) which holds when agents prefer their allocations to other allocations under all completions of the responsive set extension, and *possible completion envy-freeness* (PC EF), which holds when there exists a completion of the responsive set extension under which the fairness notion holds.

The definitions of NC EF and PC EF are underspecified in the pen-and-paper work [6]. As we will see later on, we explore two possible interpretations of these definitions, one that equates NC EF and PC EF, and the other, presented in Figure 1, enables PC EF to be trivially satisfied, leading to a simple counter-example demonstrating PC EF but not Weak SD EF. It remains open whether there are alternative interpretations of these definitions that preserve equivalence between Weak SD EF and PC EF while not equating PC EF and NC EF.

**Overview of Isabelle/HOL.** Before we delve into the formalisation, we start by introducing some Isabelle/HOL concepts and notation.

*Types and Notation.* To mechanise the setting of the assignment problem, we use Isabelle/HOL as our tool of development.

In Isabelle/HOL, each term has a type, and types in Isabelle/HOL are polymorphic, which allows us to reason about properties in multiple type classes at a time. Take the type $'a\ list$, which describes a list whose elements are of any type $'a$. We can define a function *rev* to reverse this list, and prove that $rev\,(rev\,l) = l$, which means reversing a list twice will obtain the original list. Since the property holds for a generic list, we can be sure that the property also holds for a list of numbers or a list of functions of the same type.

Isabelle/HOL's type constructors are expressive enough for our formalisation purposes. Some type constructors used in this formalisation are $'a\ set$ to describe a generic set, $'a \Rightarrow 'b$ to describe a function type with one input, $'a \Rightarrow 'b \Rightarrow 'c$ to describe a (curried) function that takes two arguments, and $'a\,relation$ to describe a predicate that takes a pair of values of type $'a$ and returns a boolean.

We use the infix notation $x \preceq [R]\, y$ to mean $R\, x\, y$, and $x \succeq [R]\, y$ to mean $R\, y\, x$.

*Locales.* In Isabelle/HOL, locales are a module system that enables describing complex theoretical structures and dependencies. Each locale defines a new context, with a fixed set of constants and/or assumptions. A theorem that holds within the context of a locale is true under the set of introduced variables and assumptions of the locale. New locales can be defined as extensions of existing locales by importing the assumptions and variables from existing locales and adding new variables or assumptions. This creates a hierarchy of locale dependencies that inherit theorems from the parent locales. We can use the Isabelle/HOL keyword *sublocale* to demonstrate locale dependency and enable the reuse of theorems proven in the context of existing locales.

Some locales we formalise and use throughout our mechanised proofs are the *random-assignment* and *discrete-assignment* locales, which describe the fractional/random assignment setting and the discrete assignment setting, respectively. In our formalisation, we also reuse locales from the Randomised Social Choice AFP entry [18]. Namely, the *pref-profile-wf* locale that defines a preference profile of agents, and the *vnm-utility* locale that introduces the von Neumann-Morgenstein utility function.

## 3 Formalising the Assignment Problem and Fairness

We start with the formalisation of the assignment problem. We follow the standard pen-and-paper definitions [6]. Figure 2 presents the Isabelle/HOL locales that underpin our formalisation of the assignment problem. These locales encode preference profiles and allocations, and the assignment problem in a modular and extensible way, supporting both the randomised (and fractional) and discrete settings.

We reuse the locale *pref-profile-wf* from the Randomised Social Choice AFP entry [18]. This locale defines a well-formed preference profile over a finite set of agents and a finite set of alternatives. Each agent is associated with a total preorder over the alternatives, representing their (possibly non-strict) ordinal preferences.

On pen-and-paper [6], an assignment $p$ is represented as an $n \times m$ matrix, where $n$ is the number of agents and $m$ is the number of alternatives, and where each entry $p[i][j]$ is the probability that alternative $j$ is assigned to agent $i$, or the fraction of $j$ assigned to $i$ in the divisible resource setting [6]. Instead, in our formalism, we weave out concepts, first defining an allocation as a function operating on a set of alternatives that maps each alternative to a real number. Then, we define an assignment $p$ that maps each agent in the set of agents to an allocation over the set of alternatives. This structure accommodates both randomised (fractional) and discrete assignments, allowing us to capture general settings in a unified formalism, separating assumptions so our theorems are more general, and with the discrete setting defined as a special case of a more general setting.

The locale *random-allocation* encodes the constraints of a valid allocation to a single agent: the set of alternatives is finite, undefined alternatives are assigned zero weight, and all values lie within the range $[0, 1]$. The *random-assignment* locale lifts this to multiple agents, ensuring that each agent receives a valid allocation. Unlike the assumption that each agent receives a lottery (i.e., a probability distribution summing to 1), we allow for general allocations where each agent's total mass is constant across all agents $\exists c.\ \forall i \in agents.\ \sum_{j \in alts} p\ i\ j = c$. This is because we are also interested in the discrete setting, and in that setting the lottery assumption has the direct consequence that the expectation of the number of objects an agent can receive is exactly 1. This generalisation

```
type_synonym (α, β) pref-profile = α ⇒ β relation

locale pref-profile-wf =
  fixes agents :: α set  and alts :: β set  and R :: (α, β) pref-profile
  assumes nonempty_agents: agents ≠ {}
  assumes nonempty_alts: alts ≠ {}
  assumes prefs_wf: ∀ i ∈ agents. finite-total-preorder-on alts (R i)
  assumes prefs_undefined: ∀ i ∉ agents. ¬ R i x y

type_synonym β allocation = β ⇒ real

locale random-allocation =
  fixes alts :: β set and  h :: β allocation
  assumes finite_alts: finite alts
  assumes undefined_alts: j ∉ alts ⟶ h j = 0
  assumes prob: ∀j ∈ alts. (0 ≤ h j ∧ h j ≤ 1)

type_synonym (α, β) assignment = α ⇒ β allocation

locale random-assignment =
  fixes agents :: α set and alts :: β set and R :: (α, β) pref-profile
  assumes pref-profile-wf agents alts R
  fixes p :: (α, β)assignment
  assumes random_alloc: ∀i ∈ agents. random-allocation alts (p i)
  assumes undefined_agent: ∀j. i ∉ agents ⟶ p i j = 0
  assumes stochastic: ∀j ∈ alts. (∑ i∈agents. p i j) = 1
  assumes sum_prob_equal_agents: ∃c :: real. ∀ i ∈ agents. (∑ j ∈ alts. p i j) = c
  assumes fin: finite agents

locale discrete_allocation = random-allocation +
  assumes disc: ∀i ∈ alts. h i = 0 ∨ h i = 1

locale discrete_assignment = random-assignment +
  assumes discr: ∀i ∈ agents. discrete_allocation alts ((p :: (α, β) assignment) i)

definition allocated-alts :: β allocation ⇒ β set ⇒ β set where
  allocated-alts p alts = {x. x ∈ alts ∧ p x ≠ 0}
```

Fig. 2: Locales: preference profiles, allocations, and assignments

permits a wider class of fair allocation scenarios while preserving compatibility with fairness criteria based on stochastic dominance. The *discrete-allocation* locale restricts *random-allocation* to binary values 0 or 1 representing full resource ownership. This constraint is propagated to the *discrete-assignment* locale that defines discrete assignments across agents. Additionally, the function *allocated-alts* allows us to interprete allocations as sets of alternatives. These definitions support reasoning about indivisible goods in the discrete setting.

Figure 2 also includes the definition of a helper function *allocated-alts*, which extracts the set of alternatives that is allocated to an agent from their allocation function. As we will see shortly, this function is essential for defining fairness notions that are based on the responsive set extension.

Our layered structure of locales and functions provides a shared foundation for stating fairness properties and verifying their relationships in both the randomised and the discrete setting. It promotes reuse and compositional reasoning in subsequent proofs about the formalisation.

**Formalizing Fairness.** We now introduce the formalisation of ten fairness notions in the assignment problem, capturing both strong and relaxed variants of two core concepts: proportionality and envy-freeness. These fairness notions are expressed using three key mathematical tools: stochastic dominance (SD), utility functions, and the responsive set extension. Figure 3 presents our formal definitions of the ten different notions of fairness, and the remainder of this section elaborates on these formal definitions. The final two definitions are ambiguous on pen-and-paper, and we discuss how we resolve the ambiguity in our formalism.

We represent all notions of fairness as a predicate over a given assignment that holds when the assignment is considered fair according to the respective fairness definition. The most prominent notions of fairness in the randomised setting are based on stochastic dominance (SD). Stochastic dominance provides a way to compare allocations when agents' preferences are represented as ordinal relations. An agent *SD-prefers* one allocation to another if, for every alternative $j$, the first allocation provides the agent with at least as many alternatives that are at least as good as $j$ as in the second allocation. We extend the classic definition of SD to allow allocations that are not necessarily lotteries, i.e., not summing to 1, by introducing a custom relation $SDA(R)$. In the definition of proportionality, the function *pmf-like-set* constructs an assignment formed of the uniform allocation (similar to a uniform distribution), where each agent receives an equal share of the alternatives.

Using this extended definition, we define the following four notions of fairness: *SD-proportional*: where each agent SD-prefers their allocation to the uniform allocation, *weak-SD-proportional*, where no agent strictly SD-prefers the uniform allocation over their own, *SD-envyfreeness*, where each agent SD-prefers their allocation to every other agent's allocation, and *weak-SD-envyfreeness*, where no agent strictly SD-prefers another agent's allocation to their own.

We also formalise fairness based on von Neumann–Morgenstern utility functions consistent with each agent's ordinal preferences. These enable defining alternative notions of fairness based on expected utility over allocations. Let

```
definition
  SDA :: β relation ⇒ (β allocation) relation where
  p ≽[SDA(R)] q ≡
    ∀x. R x x ⟶ (∑i∈{y. R y y ∧ y ≽[R] x}. p i) ≥ (∑i∈{y. R y y ∧ y ≽[R] x}. q i)

 locale vnm-utility = finite-total-preorder-on S R +
  fixes u :: 'a ⇒ real
  assumes utility_le_iff: x ∈ S ⟹ y ∈ S ⟹ u x ≤ u y ⟷ x ≼[R] y

definition RS :: β relation ⇒(β set) relation  where
  p ≽[RS(P)] q ≡ ∃ f :: (β ⇒ β). (inj_on f q ∧ f ‘ q ⊆ p ∧ (∀(x:: β)∈q. f x ≽[P] x))

definition SD-proportional :: (α, β) assignment ⇒ bool where
  SD-proportional A ≡ ∀i ∈ agents. A i ≽[SDA(R i)] pmf-like-set alts

definition weak-SD-proportional :: (α, β) assignment ⇒ bool where
  weak-SD-proportional A ≡ ∀i ∈ agents. ¬(pmf-like-set alts ≽[SDA(R i)] A i)

definition possible-proportional :: (α, β) assignment ⇒ bool where
  possible-proportional A ≡
    ∀i ∈ agents. ∃u :: (β ⇒ real).
      vnm-utility alts (R i) u ∧ sum-utility u alts (A i) ≥ sum u alts / (card agents)

definition necessary-proportional :: (α, β) assignment ⇒ bool where
  necessary-proportional A ≡
    ∀i ∈ agents. ∀u :: (β ⇒ real).
      vnm-utility alts (R i) u ⟶ sum-utility u alts (A i) ≥ sum u alts /(card agents)

definition SD-envyfreeness :: (α, β) assignment ⇒ bool where
  SD-envyfreeness A ≡ ∀i ∃ agents. ∀j ∈ agents. A i ≽[SDA(R i)] A j

definition weak_SD-envyfreeness :: (α, β) assignment ⇒ bool where
  weak_SD-envyfreeness A ≡ ∀i ∈ agents. ∀j ∈ agents. ¬ (A j ≽[SDA(R i)] A i)

definition possible-envyfreeness :: (α, β) assignment ⇒ bool where
  possible-envyfreeness A ≡
   ∀i ∈ agents. ∃u :: (β ⇒ real).
    ∀j ∈ agents. vnm-utility alts (R i) u ∧ sum-utility u alts (A i) ≥ sum-utility u alts (A j)

definition necessary-envyfreeness :: (α, β) assignment ⇒ bool where
  necessary-envyfreeness A ≡
   ∀i ∈ agents. ∀ u :: (β ⇒ real).
    ∀j ∈ agents. vnm-utility alts (R i) u ⟶ sum-utility u alts (A i) ≥ sum-utility u alts (A j)

definition possible-completion-envyfreeness :: (α, β) assignment ⇒ bool where
  possible-completion-envyfreeness A ≡
  ∀i ∈ agents. (∃P :: β set ⇒ β set ⇒ bool.
                        (∀s1 s2.  s2 ≽[RS(R i)] s1 ⟶ s2 ≽[P] s1) ∧
                        (∀j ∈ agents. allocated-alts (A i) alts ≽[P] allocated-alts (A j) alts))

definition necessary-completion-envyfreeness :: (α, β) assignment ⇒ bool where
  necessary-completion-envyfreeness A ≡
    ∀i ∈ agents. ∀P :: β set ⇒ β set ⇒ bool.
        (∀s1 s2.  s2 ≽[RS(R i)] s1 ⟶ s2 ≽[P] s1) ⟶
    (∀j ∈ agents. allocated-alts (A i) alts ≽[P] allocated-alts (A j) alts)
```

Fig. 3: Formal definitions of fairness and underlying formal definitions

*sum-utility u alts p* be defined as the sum of an agent's utility for a given allocation $p$ over alternatives *alts* using the utility function $u$, $\sum_{j \in alts}(u\ i) * (p\ i)$. The *vnm-utility* locale [18], presented in Figure 3 for convenience, provides the condition under which a utility function is a von Neumann–Morgenstern utility.

With these utility definitions, we are ready to present four more notions of fairness: *necessary-proportional*: each agent prefers their allocation at least as much as the uniform allocation using any possible utility function consistent with an agent's preferences, *possible-proportional*: there exists such a utility function for each agent, *necessary-envyfreeness*: for all consistent utility functions, agents prefer their allocation to any other agent's allocation, and *possible-envyfreeness*: for each agent, there exists a utility function under which they prefer their allocation.

In the discrete setting, we use the *responsive set extension* (RS) to lift preferences over alternatives to preferences over sets. As defined in Figure 3, given a preference relation $R$, an agent $i$ prefers set $A$ to $B$ under $RS$ if there exists an injection $f : B \to A$ such that $f(x) \succeq_i x$ for all $x \in B$.

Using this definition, we can introduce two fairness notions in the discrete setting.

We note two important clarifications regarding our formalisation of *completion-based* fairness notions. First, although our definitions are inspired by Bouveret and Lemaître [10], we allow weak (non-strict) preferences, supporting indifference, whereas their original formulation assumes strict preference. The paper we are formalising [6], refers to the original definitions [10] but is also inconsistent with the original notions in that it defines fairness under weak preferences. Second, the definition of a relation *consistent with the responsive set extension* is ambiguous in prior work. In particular, Aziz et al. [6] use this terminology but do not specify whether consistency should be interpreted as refinement (implication) or as equivalence. The original paper [10], defines a refinement $R_0$ of a relation $R$ as satisfying $R \subseteq R_0$, i.e., implication. In contrast, Aziz et al. [6] use the term *consistent* elsewhere in their article to describe equivalence (bi-implication) with respect to utility functions. To maintain compatibility with the intended refinement-based semantics of the responsive set extension, we adopt the implication interpretation: we require that the completion relation $P$ refines RS($R$), not that it is logically equivalent to it.

In fact, we first interpreted the definition with equivalence as the relation. The statement "a relation consistent with the responsive set extension of the preference relation of the agent" is formalised as $\forall s1\ s2.\ s2 \succeq [RS(R\ i)]s1 \longleftrightarrow s2 \succeq [P]s1$.

**definition** *possible-completion-envyfreeness* :: ($\alpha$, $\beta$) *assignment* $\Rightarrow$ *bool* **where**
*possible-completion-envyfreeness* A $\equiv$ $\forall$i $\in$ agents. ($\exists$P :: $\beta$ *set* $\Rightarrow$ $\beta$ *set* $\Rightarrow$ *bool*.
($\forall$s1 s2. s1 $\succeq$[$RS$(R i)] s2 $\longleftrightarrow$ s1 $\succeq$[P] s2) $\wedge$
($\forall$j $\in$ agents. *allocated-alts* (A i) alts $\succeq$[P] *allocated-alts* (A j) alts))

**definition** *necessary-completion-envyfreeness* :: ($\alpha$, $\beta$) *assignment* $\Rightarrow$ *bool* **where**
*necessary-completion-envyfreeness* A $\equiv$ $\forall$i $\in$ agents. $\forall$P :: $\beta$ *set* $\Rightarrow$ $\beta$ *set* $\Rightarrow$ *bool*.
($\forall$s1 s2. s1 $\succeq$[$RS$(R i)] s2 $\longleftrightarrow$ s1 $\succeq$[P] s2) $\longrightarrow$

($\forall$j $\in$ agents. *allocated-alts* (A i) alts $\succeq$[P] *allocated-alts* (A j) alts)

However, once the definition has been formalised, we could see that the definitions of *necessary-completion-envyfreeness* and *possible-completion-envyfreeness* would be equivalent using equivalence rather than refinement. Given an agent $i$ and preference profile $R$, the one and only $P$ relation that satisfies formula

$$\forall\ s_1\ s_2.s_1 \succeq [RS(R\ i)]s_2 \longleftrightarrow s1 \succeq [P]s2$$

is in fact the relation $RS(R\ i)$ itself. Hence, both definitions now collapse into:

$\forall$i, j $\in$ agents. *allocated-alts* (A i) alts $\succeq$[$RS$(R i)] *allocated-alts* (A j) alts

This would not be sensible if we are trying to have definitions that follow the hierarchical structure presented in Figure 1. We hence use the definitions that rely the refinement relation instead, as shown in Figure 3.

## 4 Verification

We verify relationships between the different fairness notions that we formalised, following the original pen-and-paper results. The key results of interest [6] are Theorems 2 and 3, which capture equivalences and implications among the various notions of fairness, respectively. Note that we disproved one of these results in the original paper, in particular, Equation 3b. In particular, they state the following:

$$\textit{weak-SD-proportional} \longleftrightarrow \textit{possible-proportional} \tag{1}$$

$$\textit{SD-proportional} \longleftrightarrow \textit{necessary-proportional} \tag{2}$$

$$\textit{weak-SD-envyfreeness} \longrightarrow \textit{possible-completion-envyfreeness} \tag{3a}$$

$$\textit{weak-SD-envyfreeness} \not\longleftarrow \textit{possible-completion-envyfreeness} \tag{3b}$$

$$\textit{SD-envyfreeness} \longleftrightarrow \textit{necessary-envyfreeness} \tag{4a}$$

$$\textit{necessary-envyfreeness} \longleftrightarrow \textit{necessary-completion-envyfreeness} \tag{4b}$$

$$\textit{SD-envyfreeness} \longrightarrow \textit{SD-proportional} \tag{5}$$

$$\textit{SD-proportional} \longrightarrow \textit{weak-SD-proportional} \tag{6}$$

$$\textit{possible-envyfreeness} \longrightarrow \textit{SD-proportional} \tag{7}$$

$$\textit{possible-envyfreeness} \longrightarrow \textit{weak-SD-envyfreeness} \tag{8}$$

We also prove the following simple fact in order to complete the relationships described in Figure 1:

$$\textit{necessary-envyfreeness} \longrightarrow \textit{possible-envyfreeness} \tag{9}$$

For convenience, we sometimes refer to the fairness properties to mean the property applied to a given assignment. To facilitate verification, we classify the theorems into two categories: those applicable in the randomised setting and those specific to the discrete setting. Among these eleven relationships, three involve discrete notions—namely, those based on possible and necessary completion envy-freeness. We formalise and verify these properties separately within the appropriate locales: *random-assignment* for the randomised setting and *discrete-assignment* for the discrete setting. We prove that *discrete-assignment* is a sublocale of *random-assignment*, enabling the reuse of theorems that are proven in the random setting within the discrete setting. This hierarchy ensures modularity and avoids duplicating proofs when fairness notions coincide across settings.

The fairness notions rely on the three important concepts: our extended definition of stochastic dominance, the utility function, and the responsive set extension. We divide the fairness notions into three further categories, based on which notion they refer to.

As such, once we have established the relationships between these three further notions, the equivalences and the implications among the remaining fairness notions will follow directly.

We then state our most important theorem, which lays the foundation for the relationship between fairness notions. The theorem states that the three statements below are equivalent (given two assignments $p$ and $q$, and a particular agent $i$). The paper [6] denotes the allocation of agent $i$ in assignment $p$ and $q$ as $p(i)$ and $q(i)$, respectively. They also use $\succeq_i^{SD}$ to describe the stochastic dominance relation, $\succeq_i^{RS}$ to describe the relation using the responsive set extension, both of which with regard to agent $i$'s preference profile. The set $\mathcal{U}_i$ is the set of all utility functions that are consistent with the preference relation of agent $i$. In the discrete setting, $u_i(p\,i)$ is defined as the sum of agent $i$'s valuation of the allocation [6]. However, in the context of the randomised or the fractional setting, it is more appropriate to interpret utility as a weighted sum of the agent's valuation over alternatives, where the weights correspond to the probabilities (in the random setting) or fractions (in the fractional setting) of receiving each alternative. This is the interpretation we adopt in our formalisation and verification. The equivalent statements are:

$$p\,i \succeq_i^{SD} q\,i \tag{i}$$

$$\forall u_i \in \mathcal{U}_i.\ u_i(p\,i) \geq u_i(q\,i) \tag{ii}$$

$$p\,i \succeq_i^{RS} q\,i \tag{iii}$$

In our verification effort, to prove the equivalence of these three statements, we prove the following: (i) and (ii) are equivalent, (iii) implies (ii), and (i) implies (iii). In particular, the equivalence of (i) and (ii) will be proven in the random setting, which is the locale *random-assignment*. The theorems relating (iii) to (ii), and (i) to (iii) are proven in the discrete setting (*discrete-assignment* locale). We also establish the relationship that *discrete-assignment* is a **sublocale** of *random-assignment*, which will allow us to reuse the theorem about the equivalence of (i) and (ii) in the discrete setting, so that we can claim that the three definitions are equivalent in the discrete setting.

The equivalence of (i) and (ii) is a classic result in social choice theory, and there have been both pen-and-paper proofs [11] and mechanised proofs [18]. Recall, though, that in this formalisation we generalise the notion of SD.

For the forward direction, we first prove an important fact that for each agent with a total transitive order on the set of alternatives, there exists an increasing ordering of the set of alternatives (a sequence where each element is preferred to the previous element). Given a set of alternatives, let the ordering on the alternatives be $a_1, a_2, ..., a_m$, where $a_i$ is preferred to $a_j$ for any $i \geq j$. To prove that (extended) stochastic dominance indeed implies the property of every utility function, we assume that we have two allocations $p$ and $q$ where $p \succeq [SDA(R)]q$, and $R$ is an arbitrary preference relation on the set of alternatives. We thus have $a_j \succeq [R]a_i$ for any $j \geq i$. We then show that, for any agents $i$, assuming we have any arbitrary utility function $u$ consistent with $R$ (denoting $p\,j$ and $q\,j$ as the fractions/probability of alternative $j$ in each allocations): $\sum_{j \in alts} u\,j\,(p\,j) \geq \sum_{j\ \in\ alts} u\,j\,(q\,j)$ is equivalent to $\sum_{j \in alts} u\,j\,(p\,j - q\,j) \geq 0$. The proof follows by using the Abel's summation technique [30], where summations can be rearranged as follows:

$$\sum_{j \in alts} u\,j(p\,j - q\,j) = u\,a_1(\sum_{i=1}^{m} p\,a_i - q\,a_i) + \sum_{i=2}^{m}((u\,a_i - u\,a_{i-1})(\sum_{k=i}^{m} p\,a_k - q\,a_k))$$

This is where we use our assumptions. Firstly, our original assumption for the assignment problem, where the sums of fractions in an allocation are constant across agents, makes the first summand on the right-hand side zero (note that the sequence $a_1, a_2, ..., a_m$ contains all and only the alternatives in our original set of alternatives). Regarding the second summand, we resort to the definition of our (extended) stochastic dominance to prove that each term in the summand is nonnegative, or we are proving: $\forall i \in \{2, .., m\}.(u\,a_i - u\,a_{i-1})(\sum_{k=i}^{m} p\,a_k - q\,a_k) \geq 0$ The definition can be summarised as $p \succeq [SDA(R)]q$ iff for each alternative $o$, the total fraction of alternatives that is at least as preferred as $o$ (with respect to the preference relation $R$) in $p$ is more than in $q$. For each $i$, either $i$ is the smallest index $j$ that $a_j$ is at least as preferred to as $a_i$, or there is an index $k < i$ such that $a_k \succeq [R]a_i$. The first case implies that the second factor in (1) is nonnegative, and according to the consistency of the utility function with the preference relation, the first factor is also nonnegative, hence, we must have the product in (1) nonnegative. For the second case, since there is an index $k < i$ such that $a_k \succeq [R]a_i$, $u\,a_k \geq u\,a_i$ by the consistency of utility functions. We also have that $u\,a_i \geq u\,a_k$ according to the fact that $a_i \succeq [R]a_k$ for any $i \geq k$. Thus $u\,a_k = u\,a_i$, following the fact that the utilities of all alternatives from $a_k$ leading up to $a_{i-1}$ are all equal to $u\,a_i$. As a result, the first factor is 0 in the second case, which proved our theorem in the forward direction.

For the backward direction, we assume that with any utility function consistent with the preference relation $R$, the weighted sum of utility in $p$ is greater than or equal to the weighted sum of utility in $q$. To prove that $p \succeq [SDA(R)]q$, we prove using its definition that for any alternative $o$, the sum of fractions allocated to alternatives that are at least as preferred as $o$ is larger in $p$ than in $q$. Fixing an arbitrary $x$, we prove that the sum of fractions allocated to alternatives that are at least as preferred as $x$ is larger in $p$: $\sum_{y \succeq [R] x} p\, y \geq \sum_{y \succeq [R] x} q\, y$. The utility function assumption implies $\sum_{j \in alts} u\, j\, (p\, j) \geq \sum_{j\ \in\ alts} u\, j\, (q\, j)$. We would like to pick a utility function that, when substituted in $\sum_{j \in alts} u\, j\, (p\, j)$, will give us $\sum_{y \succeq [R] x} p(y)$. One intuitive indicator function to consider is the function that returns 1 if the input alternative is strictly preferred to $x$, and returns 0 otherwise. We call this function $g$. However, since the value of the utility function is the same for all inputs preferred to $x$, this is not a von Neumann-Morgenstein utility function. We need to add a function to $g$ (possibly a utility function) to ensure it satisfies the von Neumann-Morgenstein utility function assumption. Our candidate function is another indicator-like function. This function takes an alternative $k$, and returns the number of alternatives that are at most as preferred as $k$. It is trivial to notice that this is indeed a utility function, call this function $f$, and adding $f$ to $g$ creates a utility function. However, we just want to use $g$ to reach our conclusion; thus, we want to eliminate the dependency of whether $g$ satisfies the desired condition on $f$. Notice that $g + \epsilon * f$ is a utility function for any $\epsilon > 0$. Substitute this into (2) gives us:

$$\sum_{j \in alts} (g\, j + \epsilon * f\, j)(p\, j - q\, j) \geq 0 \longleftrightarrow$$

$$\sum_{j \in alts} g\, j (p\, j - q\, j) + \epsilon * (\sum_{j \in alts} f\, j\, (p\, j - q\, j)) \geq 0$$

Knowing that $\epsilon$ can be arbitrarily small and the inequality above is true for all $\epsilon > 0$, we can prove $\sum_{j \in alts} g\, j\, (p\, j - q\, j) \geq 0$. And according to our observation of the definition of $g$, this is actually equivalent to (2). Hence, we have proven both directions of the theorem relating (i) $\Leftrightarrow$ (ii).

Moving on to the relationships: (iii) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii), in the discrete setting. Proving (iii) $\Rightarrow$ (ii) is rather trivial; all that remains is a matter of how to interpret allocations as sets of alternatives. Assuming that $p \succeq [RS(R)]q$, where $p$ and $q$ are two allocations, and $R$ is our preference relation, we want to prove that for any utility function $u$ consistent with $R$, the weighted sum of utility in $p$ is larger than that in $q$. From the definition of responsive set extension, there must be a total injective mapping $f$ from $q$ to $p$ such that $f(x) \succeq [R]x$ for any $x \in q$. With our assumption that the sum of proportions for all agents is equal, the sum of proportions in $p$ and in $q$ is equal. Since in the discrete setting, all allocations are 0 or 1, which translates to the set of alternatives with proportions being 1. As a result, the mapping $f$ above is not just injective but also bijective, which means it is a one-to-one mapping, where each element in the domain is preferred to its image under the mapping $f$. With this observation, the total utility of allocation $p$ must be larger than or equal to $q$ according to the consistency of utility functions with the preference relation

Now we prove (i) $\Rightarrow$ (iii), assuming that we have $p \succeq [SDA(R)]q$, where $p$ and $q$ are allocations which can both be either interpreted as a function or a set of alternatives, and $R$ is the preference relation, we prove that $p \succeq [RS(R)]q$. One thing to note is that, since we are also interpreting $p$ and $q$ as sets, (extended) stochastic dominance now means that for any alternative $x$, the number of alternatives that are at least as preferred to $x$ in $p$ is more than that in $q$.

To prove the responsive set property, we need to show that there exists an injection from $q$ to $p$, mapping an element in $q$ to a more preferred element in $p$. Interpreting this as a matching in a graph[6] where the set of vertices is $p \cup q$, we want to find a matching that covers all vertices in $q$.

We refer to Hall's marriage theorem to prove this result [23, 24], and we rely on a prior formalisation of this theorem in Isabelle/HOL [26, 25]. The theorem (in its graph-theoretic formulation) state that given a bipartite graph $G = (X, Y, E)$ with bipartite sets $X$ and $Y$, there is a $X$-perfect matching (a matching that covers all vertices in $X$) if and only if for all subsets $W$ of $X$, the number of elements that are adjacent to $W$ (elements in $Y$ that are adjacent to at least an element in $W$) is at least as large as the number of elements in $W$. To use this theorem, we need to prove that for any subset $t \subseteq q$, the number of elements in $p$ that are at least as preferred as some elements in $t$ is at least the cardinality of $t$. Since the preference relation is total and transitive on the set of alternatives, it is also total and transitive on any subsets of the set of alternatives, namely $t$, $p$ and $q$. We also establish that there is a minimal element $a$ in $t$. Since $a$ is the minimal element in $t$, and $t \subseteq q$, it is a subset of the set containing all elements that are at least as preferred to $a$ in $q$. This subset then has smaller cardinality than the set containing all elements that are at least as preferred to $a$ in $p$, according to (extended) stochastic dominance. The set containing all elements in $p$ that are at least as preferred to $a$, is then a subset of the set containing all elements in $p$ that are at least as preferred to an element in $t$. As a result, we must have that the number of elements in $p$ that is at least as preferred to an element in $t$ is at least as large as the number of elements in $t$, hence there exists such a matching. In our proof, the observation is that the matching is the relation $\succeq [R]$, where an element (alternative) $a$ is adjacent to an element $b$ iff $a \succeq [R] b$. Since (i) $\Leftrightarrow$ (ii), (iii) $\Rightarrow$ (ii), and (i) $\Rightarrow$ (iii), we can conclude that the three statements are equivalent in the discrete setting.

One more important result remains before we investigate relationships between fairness notions. We prove that a *weaker* version of (extended) stochastic dominance would be equivalent to a weakened notion for utility functions. We prove that an allocation $q$ not strictly *SD preferred* to another allocation $p$ is equivalent to the fact that there exists a utility function that makes the weighted sum of the utility of alternatives in $p$ is larger than or equal to the weighted sum of the utility of alternatives in $q$. Here, $q$ is strictly SD preferred to $p$ means $q \succeq [SDA(R)] p \wedge \neg p \succeq [SDA(R)] q$. What we need to prove is:

$$\neg p \succeq [SDA(R)] q \longleftrightarrow \exists u. u \in \mathcal{U} \wedge \sum_{j \in alts} (u\, j)\, (p\, j) \geq \sum_{j \in alts} (u\, j)\, (q\, j)$$

Here, $\mathcal{U}$ is the set of all utility functions that are consistent with the preference relation $R$. This equivalence is already formalised and proved in the work [18], our proof adopts this proof, with some minor modifications taking into account our weakened assumption that the allocations are not necessarily lotteries.

**Relationships between fairness notions.** With our new definition in contrast to the original definition by [6], a modification to the equivalence is needed, where instead of an Equivalence (3a, 3b) as per the pen-and-paper proof, only Implication (3a) from *weak-SD-envyfreeness* to *possible-completion-envyfreeness* holds. We have a counterexample showing that *possible-completion-envyfreeness* does not imply *weak-SD-envyfreeness* that is formalised in Isabelle/HOL. It remains to be seen whether an alternative definition of *possible-completion-envyfreeness* implies *weak-SD-envyfreeness* while retaining the current relationships with other notions of fairness.

With the equivalence in the previous section about stochastic dominance and utility functions, Equivalence (2) and the equivalence of *SD-envyfreeness* and *necessary-envyfreeness* in Equivalence (4a) conveniently follow. The equivalence of *necessary-envyfreeness* and *necessary-completion-envyfreeness* in Equivalence (4b), and Implication (3a) follow from the equivalence between the stochastic dominance relation and the responsive set extension relation. Equivalence (1) follows from the equivalence of (i) and (ii).

The implication of *possible-envyfreeness* for *possible-proportionality* and the implication of *necessary-envy-freeness* for *necessary-proportionality* follows directly from the observation that if an agents value their allocations at least as much as other agents, then they must value their allocations at least as much as the equal share allocation (or average share allocation). Combine this with Equivalences (2), (4a), (4b), we should obtain Implications (5) and (7). Implication (6) is rather trivial, as the latter definition is a weaker variant of the former one. Implications (8) also follows from the equivalence of (i) and (ii).

## 5 Related Work

The formalisation and verification of foundational results in social choice theory using proof assistants has received increasing attention in the formal methods community. Much of the existing work focuses on voting systems, particularly formalising classical impossibility theorems and properties of collective decision-making rules or on certifying voting mechanisms.

Among the most well-known results, Arrow's impossibility theorem [2] has been formalised multiple times. Notably, Nipkow [32, 31] and Gammie [20] each provided formalisations in Isabelle/HOL, while Wiedijk formalised the theorem in Mizar [41]. Gammie's formalisation further extends to incorporate May's majority decision theorem [28] and Sen's liberal paradox [40]. May's theorem has also been formalised independently by Li [27] in the Roqc Prover (formerly Coq). Nipkow's Isabelle/HOL formalisation [32] also includes a proof of the Gibbard–Satterthwaite theorem [22, 39], another cornerstone result in voting theory. While these efforts have significantly contributed to the reliable formalisation of classical results in social choice theory, their primary focus is on the aggregation of preferences to select a winning candidate or to generate a collective ranking. In contrast, our work addresses a distinct branch of social choice theory: the fair allocation of indivisible goods, which focuses on distributing resources across individuals under ordinal preferences.

In the domain of resource allocation, there has been some foundational formalisation work. Parsert and Kaliszyk [36] as well as Eberl [18] have independently formalised the von Neumann–Morgenstern (vNM) utility function in Isabelle/HOL. Eberl's work on formalising randomised social choice further includes the notions of preference profiles, lottery-based social choice functions, and a notion of stochastic dominance for random assignments. These formalised concepts provide a critical basis for reasoning about randomised outcomes in collective decision-making.

Our work builds directly on Eberl's formalisation, reusing the definitions of preference profiles and utility function to formalise the assignment problem setting, with a generalisation of Eberl's stochastic dominance definition to account for the discrete setting of the assignment problem.

This is the first formalisation of the assignment problem along with associated fairness notions such as SD-proportionality and SD-envy-freeness. The formalisation verifies known relationships between these notions and formally refutes one pen-and-paper relationship that was thought to be true [6]. Similar to the pen-and-paper proofs, our proofs rely on Hall's theorem [23]. Our formalisation relies on an existing Isabelle/HOL formalisation of this theorem [26].

## 6 Conclusion

We presented a formalisation of the assignment problem in Isabelle/HOL, covering both the fractional (randomised) and discrete settings. Our formalisation defines the core structure of assignment problems, introduces multiple notions of fairness across these settings, and provides formal proofs about how they relate.

## References


1. Amanatidis, G., Aziz, H., Birmpas, G., Filos-Ratsikas, A., Li, B., Moulin, H., Voudouris, A.A., Wu, X.: Fair division of indivisible goods: Recent progress and open questions. Artif. Intell. **322**(C) (Sep 2023). https://doi.org/10.1016/j.artint.2023.103965, `https://doi.org/10.1016/j.artint.2023.103965`
2. Arrow, K.J.: Social Choice and Individual Values. Yale University Press (2012), `http://www.jstor.org/stable/j.ctt1nqb90`
3. Aziz, H.: Developments in multi-agent fair allocation. In: The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020. pp. 13563–13568. AAAI Press (2020). https://doi.org/10.1609/AAAI.V34I09.7082, `https://doi.org/10.1609/aaai.v34i09.7082`
4. Aziz, H., Brandt, F., Elkind, E., Skowron, P.: Computational social choice: The first ten years and beyond. In: Steffen, B., Woeginger, G.J. (eds.) Computing and

Software Science - State of the Art and Perspectives, Lecture Notes in Computer Science, vol. 10000, pp. 48–65. Springer (2019). https://doi.org/10.1007/978-3-319-91908-9_4, https://doi.org/10.1007/978-3-319-91908-9_4

5. Aziz, H., Brandt, F., Stursberg, P.: On popular random assignments. In: Vöcking, B. (ed.) Algorithmic Game Theory. pp. 183–194. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
6. Aziz, H., Gaspers, S., Mackenzie, S., Walsh, T.: Fair assignment of indivisible objects under ordinal preferences. Artificial Intelligence **227**, 71–92 (2015). https://doi.org/https://doi.org/10.1016/j.artint.2015.06.002, https://www.sciencedirect.com/science/article/pii/S0004370215000880
7. Aziz, H., Luo, P., Rizkallah, C.: Incompatibility of efficiency and strategyproofness in the random assignment setting with indifferences. Economics Letters **160**, 46–49 (2017). https://doi.org/https://doi.org/10.1016/j.econlet.2017.08.010, https://www.sciencedirect.com/science/article/pii/S0165176517303336
8. Aziz, H., Luo, P., Rizkallah, C.: Rank maximal equal contribution: A probabilistic social choice function. In: McIlraith, S.A., Weinberger, K.Q. (eds.) Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. pp. 910–916. AAAI Press (2018). https://doi.org/10.1609/AAAI.V32I1.11448, https://doi.org/10.1609/aaai.v32i1.11448
9. Bogomolnaia, A., Moulin, H.: A new solution to the random assignment problem. J. Econ. Theory **100**(2), 295–328 (2001). https://doi.org/10.1006/JETH.2000.2710, https://doi.org/10.1006/jeth.2000.2710
10. Bouveret, S., Endriss, U., Lang, J.: Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In: Coelho, H., Studer, R., Wooldridge, M.J. (eds.) ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings. Frontiers in Artificial Intelligence and Applications, vol. 215, pp. 387–392. IOS Press (2010). https://doi.org/10.3233/978-1-60750-606-5-387, https://doi.org/10.3233/978-1-60750-606-5-387
11. Brandl, F., Brandt, F., Eberl, M., Geist, C.: Proving the incompatibility of efficiency and strategyproofness via SMT solving. J. ACM **65**(2), 6:1–6:28 (2018). https://doi.org/10.1145/3125642, https://doi.org/10.1145/3125642
12. Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D.: Handbook of Computational Social Choice. Cambridge University Press, USA, 1st edn. (2016)
13. Brandt, F., Geist, C.: Finding strategyproof social choice functions via SAT solving. J. Artif. Intell. Res. **55**, 565–602 (2016). https://doi.org/10.1613/JAIR.4959, https://doi.org/10.1613/jair.4959
14. Brandt, F., Geist, C., Peters, D.: Optimal bounds for the no-show paradox via SAT solving. Math. Soc. Sci. **90**, 18–27 (2017). https://doi.org/10.1016/J.MATHSOCSCI.2016.09.003, https://doi.org/10.1016/j.mathsocsci.2016.09.003
15. Caragiannis, I., Gravin, N., Huang, X.: Envy-freeness up to any item with high nash welfare: The virtue of donating items. In: Karlin, A.R., Immorlica, N., Johari, R. (eds.) Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019. pp. 527–545. ACM (2019). https://doi.org/10.1145/3328526.3329574, https://doi.org/10.1145/3328526.3329574

16. Chakraborty, M., Segal-Halevi, E., Suksompong, W.: Weighted fairness notions for indivisible items revisited. ACM Transactions on Economics and Computation **12**(3) (Sep 2024). https://doi.org/10.1145/3665799, https://doi.org/10.1145/3665799
17. Do, D.M., Rizkallah, C.: Formalising fairness in the assignment problem with ordinal preferences in isabelle/hol (artefact), https://github.com/crizkallah/crizkallah.github.io/blob/master/formalisations/Fairness_Fractional_Assignment.thy, last accessed July 2025
18. Eberl, M.: Randomised social choice theory. Archive of Formal Proofs (May 2016), https://isa-afp.org/entries/Randomised_Social_Choice.html, Formal proof development
19. Endriss, U.: Lecture notes on fair division. CoRR **abs/1806.04234** (2018), http://arxiv.org/abs/1806.04234
20. Gammie, P.: Some classical results in social choice theory. Archive of Formal Proofs (November 2008), https://isa-afp.org/entries/SenSocialChoice.html, Formal proof development
21. Gehrlein, W.V.: Condorcet's paradox. Theory and Decision **15**(2), 161–197 (1983). https://doi.org/10.1007/BF00143070, https://doi.org/10.1007/BF00143070
22. Gibbard, A.: Manipulation of voting schemes: A general result. Econometrica **41**(4), 587–601 (1973), http://www.jstor.org/stable/1914083
23. Hall, P.: On representatives of subsets. Classic Papers in Combinatorics pp. 58–62 (1987)
24. Halmos, P.R., Vaughan, H.E.: The marriage problem. In: Classic Papers in Combinatorics, pp. 146–147. Springer (2009)
25. Jiang, D., Nipkow, T.: Hall's marriage theorem. Archive of Formal Proofs (December 2010), https://isa-afp.org/entries/Marriage.html, Formal proof development
26. Jiang, D., Nipkow, T.: Proof pearl: The marriage theorem. In: Jouannaud, J., Shao, Z. (eds.) Proceedings of the First International Conference on Certified Programs and Proofs CPP 2011, Kenting, Taiwan, December 7-9, 2011. Lecture Notes in Computer Science, vol. 7086, pp. 394–399. Springer (2011). https://doi.org/10.1007/978-3-642-25379-9_28, https://doi.org/10.1007/978-3-642-25379-9_28
27. Li, K.H.: Formalizing May's theorem. CoRR **abs/2210.05342** (2022). https://doi.org/10.48550/ARXIV.2210.05342, https://doi.org/10.48550/arXiv.2210.05342
28. May, K.O.: A set of independent necessary and sufficient conditions for simple majority decision. Econometrica **20**(4), 680–684 (1952). https://doi.org/10.2307/1907651
29. Mishra, S., Padala, M., Gujar, S.: Fair allocation of goods and chores - tutorial and survey of recent results. CoRR **abs/2307.10985** (2023). https://doi.org/10.48550/ARXIV.2307.10985, https://doi.org/10.48550/arXiv.2307.10985
30. Niculescu, C.P., Stănescu, M.M.: A note on abel's partial summation formula. Aequationes mathematicae **91**, 1009–1024 (2017)
31. Nipkow, T.: Arrow and Gibbard-Satterthwaite. Archive of Formal Proofs (September 2008), https://isa-afp.org/entries/ArrowImpossibilityGS.html, Formal proof development
32. Nipkow, T.: Social choice theory in HOL. Journal of Automated Reasoning **43**(3), 289–304 (2009). https://doi.org/10.1007/S10817-009-9147-4, https://doi.org/10.1007/s10817-009-9147-4

33. Pacuit, E.: Voting Methods. In: Zalta, E.N., Nodelman, U. (eds.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Summer 2024 edn. (2024)
34. Parsert, J., Kaliszyk, C.: Microeconomics and the first welfare theorem. Archive of Formal Proofs (September 2017), `https://isa-afp.org/entries/First_Welfare_Theorem.html`, Formal proof development
35. Parsert, J., Kaliszyk, C.: Towards formal foundations for game theory. In: Avigad, J., Mahboubi, A. (eds.) Interactive Theorem Proving (ITP). Lecture Notes in Computer Science, vol. 10895, pp. 495–503. Springer (2018). https://doi.org/10.1007/978-3-319-94821-8_29, `https://doi.org/10.1007/978-3-319-94821-8\_29`
36. Parsert, J., Kaliszyk, C.: Von-neumann-morgenstern utility theorem. Archive of Formal Proofs (July 2018), `https://isa-afp.org/entries/Neumann_Morgenstern_Utility.html`, Formal proof development
37. Pruhs, K., Woeginger, G.J.: Divorcing made easy. In: Kranakis, E., Krizanc, D., Luccio, F.L. (eds.) Fun with Algorithms - 6th International Conference, FUN 2012, Venice, Italy, June 4-6, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7288, pp. 305–314. Springer (2012). https://doi.org/10.1007/978-3-642-30347-0_30, `https://doi.org/10.1007/978-3-642-30347-0\_30`
38. Satterthwaite, M.A.: Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory **10**(2), 187–217 (1975). https://doi.org/https://doi.org/10.1016/0022-0531(75)90050-2, `https://www.sciencedirect.com/science/article/pii/0022053175900502`
39. Satterthwaite, M.A.: Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory **10**(2), 187–217 (1975). https://doi.org/https://doi.org/10.1016/0022-0531(75)90050-2, `https://www.sciencedirect.com/science/article/pii/0022053175900502`
40. Sen, A.K.: The impossibility of a paretian liberal. Journal of Political Economy **78**(1), 152–157 (1970)
41. Wiedijk, F.: Formalizing arrow's theorem. Sadhana **34**(1), 193–220 (2009)