The possibility of judgment aggregation on agendas with subjunctive implications

Franz Dietrich[1]
April 2005, revised May 2007
to appear in *Journal of Economic Theory*

Abstract. The new field of judgment aggregation aims to find collective judgments on logically interconnected propositions. Recent impossibility results establish limitations on the possibility to vote independently on the propositions. I show that, fortunately, the impossibility results do not apply to a wide class of realistic agendas once propositions like "if a then b" are adequately modelled, namely as subjunctive implications rather than material implications. For these agendas, consistent and complete collective judgments can be reached through appropriate quota rules (which decide propositions using acceptance thresholds). I characterise the class of these quota rules. I also prove an abstract result that characterises consistent aggregation for arbitrary agendas in a general logic.

Key words: judgment aggregation, subjunctive implication, material implication, characterisation of possibility agendas

JEL Classification Numbers: D70, D71, D79

1 Introduction

In judgment aggregation, the objects of the group decision are not as usual (mutually exclusive) *alternatives*, but *propositions* representing interrelated (yes/no) questions the group faces. To ensure that these interrelations are well-defined, propositions are statements in a formal logic. As a simple example, suppose the three-member board of a central bank disagrees on which of the following propositions hold.

a :	GDP growth will pick up.
b :	Inflation will pick up.
$a \to b$:	*If* GDP growth will pick up *then* inflation will pick up.

Reaching collective beliefs is non-trivial. In Table 1, each board member holds consistent (yes/no) beliefs but the propositionwise majority beliefs are inconsistent. To achieve consistent collective judgments, the group cannot use majority voting. What procedure should the group use instead? A wide-spread view is that, in this

[1]This paper was presented at the *Risk, Uncertainty and Decision* seminar (MSE, Paris, April 2005), the *Aggregation of Opinions* workshop (Yale Law School, September 2006), and the 8th *Augustus de Morgan* workshop (King's College London, November 2006). Benjamin Polak's extensive comments and suggestions have benefited the paper on substantive and presentational levels. I am also grateful for helpful comments by two referees and by Christian List and Philippe Mongin.

	a	$a \to b$	b
1/3 of the board	Yes	Yes	Yes
1/3 of the board	No	Yes	No
1/3 of the board	Yes	No	No
Collective under majority rule	Yes	Yes	No
Collective under premise-based rule	Yes	Yes	Yes
Collective under the (below-defined) quota rule	No	No	No

Table 1: A simple judgment aggregation problem and three aggregation rules

as in most other judgment aggregation problems, we must give up aggregating propositionwise[2], for instance in favour of a premise-base rule (as discussed below). I show that this conclusion is often an artifact of an inappropriate way to model implications like $a \to b$. In many judgment aggregation problems, a more appropriate *subjunctive* interpretation of implications changes the logical relations between propositions in such a way that we can aggregate on a propositionwise basis without creating collective inconsistencies. Indeed, we can use *quota rules*: here, separate anonymous votes are taken on each proposition using (proposition-specific) acceptance thresholds. Suppose for instance that the following thresholds are used: b is accepted if and only if a majority accepts b, and a $p \in \{a, a \to b\}$ is accepted if and only if at least 3/4 of people accept p. Then, in the situation of Table 1, a, $a \to b$ and b are all rejected, i.e. the outcome is $\{\neg a, \neg(a \to b), \neg b\}$.

The problem is that this outcome, although intuitively perfectly consistent, is declared *in*consistent in classical logic, because classical logic defines $\neg(a \to b)$ as equivalent to $a \wedge \neg b$ ("a and not-b"), by interpreting "$\to$" as a *material* rather than a *subjunctive* implication. Is this equivalence plausible in our example? Intuitively, $a \wedge \neg b$ does indeed entail $\neg(a \to b)$, but $\neg(a \to b)$ does not entail $a \wedge \neg b$ because $\neg(a \to b)$ does not intend to say anything about whether a and b are *actually* true or false: rather it intends to say that b *would* be false in the *hypothetical* (hence possibly counterfactual) case of a's truth. Indeed, a person who believes that it is false that a pick up in GDP growth leads to a pick up in inflation may or may not believe that GDP growth or inflation will *actually* pick up; what he believes is rather that inflation will not pick up in the *hypothetical* case(s) that GDP growth will pick up.

In real-life judgment aggregation problems, implication statements usually have a subjunctive meaning. It is important not to misrepresent this meaning using material implications and classical logic, because this creates unnatural logical connections and artificial impossibilities of aggregation. The above quota rule, for instance, guarantees collective consistency (not just for the profile in Table 1) if the implication "$a \to b$" is subjunctive, but not if it is material. More generally, I establish the existence of quota rules with consistent outcomes for a large class of realistic agendas: the so-called *implication agendas*, which contain (bi-)implications and atomic propositions. This possibility is created by interpreting (bi-)implications subjunctively; it disappears if we instead use classical logic, i.e. interpret (bi-)implications materially. At first sight, this positive finding seems in conflict with the recent surge of impossibility results

[2] That is, aggregating by voting independently on the propositions: the collective judgment on any proposition p depends only on how the individuals judge p, not on how they judge other propositions. This property is usually called "independence".

on propositionwise aggregation (see below). In fact, these results presuppose logical interconnections between propositions that are stronger than (or different to) those which I obtain here under the subjunctive interpretation of (bi-)implications. In various results, I derive the (necessary and sufficient) conditions that the acceptance thresholds of quota rules must satisfy in order to guarantee consistent outcomes. These results are applications of an abstract characterisation result, Theorem 3, which is valid for arbitrary agendas in a general logic. It also generalises the "intersection property" result by Nehring and Puppe [18, 19] (but not that by Dietrich and List [6]).

Although I show that collective consistency is often achievable by aggregating propositionwise (using quota rules), I do not wish to generally advocate propositionwise aggregation. In particular, one may reject propositionwise aggregation rules by arguing that they neglect *relevant information*: in order to decide on b it is arguably not just relevant how people judge b but also *why* they do so, i.e. how they judge b's "premises" a and $a \rightarrow b$. This naturally leads to the popular *premise-based* rule: here, only a and $a \rightarrow b$ – the "premises" – are decided through (majority) votes, while b – the "conclusion" – is accepted if and only if a and $a \rightarrow b$ have been accepted; so that, in the situation of Table 1, a and $a \rightarrow b$, and hence b, are accepted.

Despite the mentioned objection, propositionwise aggregation rules are superior from a manipulation angle: non-propositionwise aggregation rules can be manipulated by agenda setters (Dietrich [2]) and by voters (Dietrich and List [5]).[3]

In general, the judgment aggregation problem – deciding which propositions to accept based on which ones the individuals accept – and its formal results are open to different interpretations of "accepting" and different sorts of propositions. This paper's examples and discussion focus on the case that "accepting" means "believing",[4] and mostly on the case that the propositions have a *descriptive* content (like "GDP growth will pick up"), although Section 4 touches on *normative* propositions (like "peace is better than war").[5]

In the literature, judgment aggregation is discussed on a less formal basis in law (e.g. Kornhauser and Sager [12], Chapman [1]) and political philosophy (e.g. Pettit [22]), and is formalised in List and Pettit [15] who use classical propositional logic. Also the related *belief merging* literature in artificial intelligence uses classical propositional logic to represent propositions (e.g. Konieczny and Pino-Perez [11] and Pigozzi [23]). A series of results establish, for different agendas, the impossibility of aggregating on a propositionwise basis in accordance with collective consistency and different other conditions (e.g. List and Pettit [15], Pauly and van Hees [21], Dietrich [2, 4], Gärdenfors [10], Nehring and Puppe [20], van Hees [26], Dietrich and List [7], Dokow and Holzman [9] and Mongin [17]). Further impossibilities (with minimal

[3]Consider for instance premise-based voting in Table 1. The agenda setter may reverse the outcome on b by replacing the premises a and $a \rightarrow b$ by other premises a' and $a' \rightarrow b$. Voter 2 or 3 can reverse the outcome on b by pretending to reject *both* premises a and $a \rightarrow b$.

[4]Judgment aggregation is the aggregation of belief sets if "accepting" means "believing", the aggregation of desire sets if "accepting" means "desiring", the aggregation of moral judgment sets if "accepting" means "considering as morally good", etc.

[5]By considering beliefs on possibly normative propositions, judgment aggregation uses a broader "belief" notion than is common in economics, where beliefs usually apply to descriptive facts only. For instance, standard preference aggregation problems can be modelled as judgment aggregation problems by interpreting preferences as *beliefs* of normative propositions like "x is better than y" (see Dietrich and List [7]; also List and Pettit [16]).

agenda conditions) follow from Nehring and Puppe's [18, 19] results on strategy-proof social choice. To achieve possibility, propositionwise aggregation is given up in favour of *distance-based* aggregation by Pigozzi [23] (drawing on Konieczny and Pino-Perez [11]), of *sequential* aggregation by List [14] and Dietrich and List [6], and of aggregating *relevant information* by Dietrich [4].

Section 7 uses Dietrich's [3] judgment aggregation model in general logics, and the other sections use for the first time possible-worlds semantics.

2 Definitions

We consider a group or persons $N = \{1, 2, \ldots, n\}$ ($n \geq 2$), who need collective judgments on a set of propositions expressed in formal logic.

The language. Following Dietrich's [3] general logics model, a language is given by a non-empty set $\mathbf{L}$ of sentences (called *propositions*) closed under negation, i.e. $p \in \mathbf{L}$ implies $\neg p \in \mathbf{L}$. (Interesting languages of course have also other connectives than negation $\neg$). Logical interconnections are captured either by an *entailment relation* $\vDash$ (telling for which $A \subseteq \mathbf{L}$ and $p \in \mathbf{L}$ we have $A \vDash p$) or, equivalently, by a *consistency* notion (telling which sets $A \subseteq \mathbf{L}$ are consistent).[6] The two notions are interdefinable: a set $A \subseteq \mathbf{L}$ is inconsistent if and only $A \vDash p$ and $A \vDash \neg p$ for some $p \in \mathbf{L}$; and an entailment $A \vDash p$ holds if and only if $A \cup \{\neg p\}$ is inconsistent.[7] The precise nature of logical interconnections is addressed later. A proposition $p \in \mathbf{L}$ is a *contradiction* if $\{p\}$ is inconsistent, and a *tautology* if $\{\neg p\}$ is inconsistent.

All following sections except Section 7 consider a *particular* language: $\mathbf{L}$ is the set of propositions constructible using $\neg$ ("not"), $\wedge$ ("and") and $\rightarrow$ ("if-then") from a set $\mathcal{A} \neq \emptyset$ of non-decomposable symbols, called *atomic* propositions (and representing simple statements like "inflation will pick up"). So $\mathbf{L}$ is the smallest set such that (i) $\mathcal{A} \subseteq \mathbf{L}$ and (ii) $p, q \in \mathbf{L}$ implies $\neg p \in \mathbf{L}$, $(p \wedge q) \in \mathbf{L}$ and $(p \rightarrow q) \in \mathbf{L}$. The critical question, treated in the next section, is how (not) to define the logical interconnections on $\mathbf{L}$: while some entailments like $a, b \vDash a \wedge b$ and $a, a \rightarrow b \vDash b$ are not controversial, others are. Notationally, I drop brackets when there is no ambiguity, e.g. $c \rightarrow (a \wedge b)$ stands for $(c \rightarrow (a \wedge b))$. Further, $p \vee q$ ("p or q") stands for $\neg(\neg p \wedge \neg q)$, and $p \leftrightarrow q$ ("p if and only if q") stands for $(p \rightarrow q) \wedge (q \rightarrow p)$. For any conjunction $p = a_1 \wedge ... \wedge a_k$ of one or more atomic propositions $a_1, ..., a_k$ (called the *conjuncts* of p), let $\mathrm{C}(p) := \{a_1, ..., a_k\}$ (e.g. $\mathrm{C}(a) = \{a\}$ and $\mathrm{C}(a \wedge b) = \mathrm{C}(b \wedge a) = \{a, b\}$).

In judgment aggregation, the term "connection rule" commonly refers to implicational statements like "*if* GDP growth continues *and* interest rates stay below X *then* inflation will rise". I now formalise this terminology. If each of p and q is a conjunction of one or more atomic propositions,

[6] For the two approaches, see Dietrich [3]. Logical interconnections can be interpreted either *semantically* or *syntactically* (in the latter case, the symbol "$\vdash$" is more common than "$\vDash$"). In the (classical or non-classical) logics considered in Sections 3-6, I define interconnections semantically (but there are equivalent syntactic definitions). Dropping brackets, I often write $p_1, ..., p_k \vDash p$ for $\{p_1, ..., p_k\} \vDash p$.

[7] The latter equivalence supposes that the logic is not paraconsistent. All logics considered in this paper are of this kind.

- $p \to q$ is a *uni-directional connection rule*, called *non-degenerate* if $C(q) \backslash C(p) \neq \emptyset$, i.e. if $p \to q$ is not a tautology (under the classical or the non-classical entailment relation discussed later);
- $p \leftrightarrow q$ is a *bi-directional connection rule*, called *non-degenerate* if $C(q) \backslash C(p) \neq \emptyset$ and $C(p) \backslash C(q) \neq \emptyset$, i.e. if neither $p \to q$ nor $q \to p$ is a tautology.

A uni- or bi-directional connection rule is simply called a *connection rule*.

The agenda. The *agenda* is the set of propositions on which decisions are needed. Formally, it is a non-empty set $X \subseteq \mathbf{L}$ of the form $X = \{p, \neg p : p \in X^+\}$ for some set X^+ containing no negated proposition $\neg q$. In the introductory example, $X^+ = \{a, b, a \to b\}$. Notationally, double-negations cancel each other out: if $p \in X$ is a negated proposition $\neg q$ then hereafter when I write "$\neg p$" I mean q rather than $\neg\neg q$. (This ensures that $\neg p \in X$.)

An agenda$X$ (in the language $\mathbf{L}$ just defined) is an *implication* agenda if X^+ consists of non-degenerate connection rules and the atomic propositions occurring in them; it is called *simple* if all its connection rules are uni-directional ones $p \to q$ in which p and q are atomic propositions.

Many standard examples of judgment aggregation problems can be modelled with implication agendas. The atomic propositions represent (controversial) issues, and connection rules represent (controversial) links between issues. Any accepted connection rule establishes a constraint on how to decide the issues.

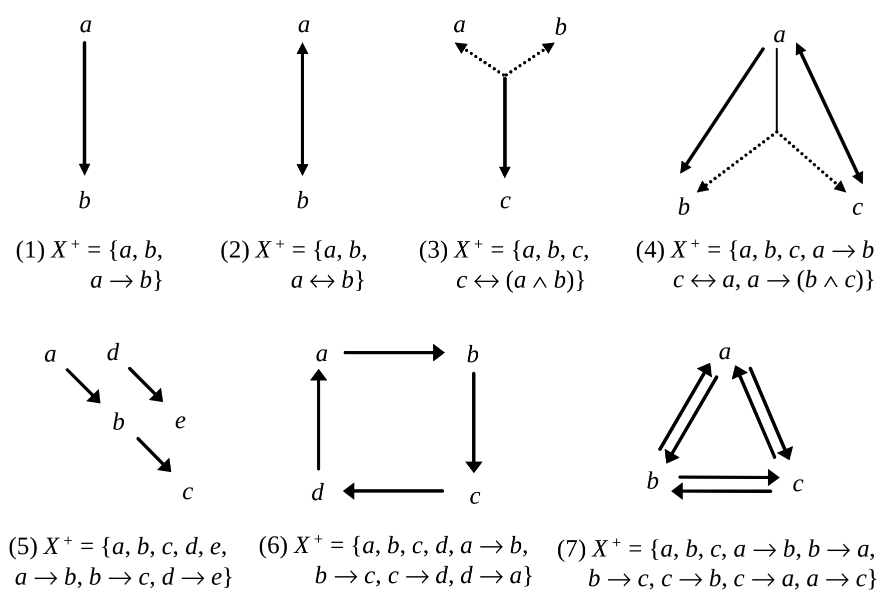

Figure 1: Seven implication agendas, represented as networks.

Implication agendas can always be represented graphically as networks over its atomic propositions.[8] Figure 1 shows seven implication agendas, of which (1) and

[8]Nodes contain atomic propositions. Arrows represent connection rules: bi-directional arrows indicate bi-implications, and bifurcations indicate conjunctions of more than one atomic proposition.

(5)-(7) are simple. Agenda (1) represents our central bank example. An environmental expert commission might face the agenda (2), where a is "global warming will continue" and b is "the ozone hole exceeds size X". The judges of a legal court face, in the decision problem from which judgment aggregation originated, an agenda of type (3), where a is "the defendant has broken the contract", b is "the contract is legally valid", c is "the defendant is liable", and $c \leftrightarrow (a \wedge b)$ is a claim on what constitutes (necessary and sufficient) conditions of liability.[9] A company board trying to predict the price policy of three rival firms A-C might face the agenda (3) or (4) or (7), where a is "Firm A will raise prices", b is "Firm B will raise prices", and c is "Firm C will raise prices". The three agendas differ in the type of connections between a, b, c deemed possible.

In Section 4, I discuss two types of decision problems captured by implication agendas: reaching judgments on facts and their causal relations, and reaching judgments on hypotheses and their justificational/evidential relations.

But not all realistic judgment aggregation problems are formalisable by implication agendas. Some judgment aggregation problems involve a generalised kind of implication agenda, obtained by generalising the definition of connection rules so as to include (bi-)implications between propositions p and q other than conjunctions of atomic propositions.[10] More radical departures from implication agendas include: (i) the agenda given by $X^+ = \{a, b, a \wedge b\}$, which contains no connection rule but the Boolean expression $a \wedge b$; (ii) the agenda representing a preference aggregation problem, which contains propositions of the form xRy from a predicate logic (see Section 7); (iii) agendas where X^+ contains only atomic propositions, between which certain connection rules are imposed exogenously (rather than subjected to a decision).

Judgment sets. A *judgment set* (held by a person or the group) is a subset $A \subseteq X$; $p \in A$ stands for "the person/group accepts proposition p". A judgment set A can be more or less rational. Ideally, it should be both *complete*, i.e. contain at least one member of each pair $p, \neg p \in X$, and *(logically) consistent*. A is *weakly consistent* if A does not contain a pair $p, \neg p \in X$ (i.e., intuitively, if A is not "obviously inconsistent"). For agenda (1) in Figure 1, $\{a, a \to b, \neg b\}$ is complete, weakly consistent, but not consistent because A entails b (in fact, $\{a, a \to b\}$ entails b) and A entails $\neg b$ (in fact, contains $\neg b$). So to say, "weak consistency" means not to *contain* a contradiction $p, \neg p$, and consistency means not to *entail* one.

Aggregation rules. A *profile* is an n-tuple $(A_1, \ldots, A_n)$ of (individual) judgment sets $A_i \subseteq X$. A *(judgment) aggregation rule* is a function F that maps each profile $(A_1, \ldots, A_n)$ from a given domain of profiles to a (group) judgment set $F(A_1, \ldots, A_n) = A \subseteq X$. The domain of F is *universal* if it consists of all profiles of complete and consistent judgment sets. F is *complete/consistent/weakly consistent* if F generates a complete/consistent/weakly consistent judgment set for each profile in

[9] A *doctrinal paradox* arises if there is a majority for a, a majority for b, a unanimity for $c \leftrightarrow (a \wedge b)$, but a majorty for $\neg c$.

[10] If, for instance, p and q were allowed to be *dis*junctions of atomic propositions then $a \to (b \vee c)$ would count as a connection rule, so that $X^+ = \{a, b, c, (a \vee b) \to c\}$ would define an implication agenda of the so-generalised kind. Generalised implication agendas may well be relevant as groups may need to make up their mind on generalised types of connection rules. The possibility of consistent aggregation by quota rules may disappear for such agendas. So our subjunctive reading of (bi-)implications (see Section 3) is not a general recipe for possibility in judgment aggregation.

its domain. On the universal domain, *majority rule* (given by $F(A_1, ..., A_n) = \{p \in X : \text{more persons } i \text{ have } p \in A_i \text{ than } p \notin A_i\}$) is weakly consistent, and a *dictatorial rule* (given by $F(A_1, ..., A_n) = A_j$ for a fixed j) is even consistent. We will focus on *quota rules* thus defined. To each family $(m_p)_{p \in X^+}$ of numbers in $\{1, ..., n\}$, the *quota rule with thresholds* $(m_p)_{p \in X^+}$ is the aggregation rule with universal domain given by

$$F_{(m_p)_{p \in X^+}}(A_1, ..., A_n) = \{p \in X : \text{at least } m_p \text{ persons } i \text{ have } p \in A_i\},$$

where $m_{\neg p} := n - m_p + 1$ for all $p \in X^+$ to ensure exactly one member of each pair $p, \neg p \in X$ is accepted, i.e. that quota rules are complete and weakly consistent.

So each family of thresholds $(m_p)_{p \in X^+}$ in $\{1, ..., n\}$ generates a quota rule. As one easily checks, an aggregation rule is a quota rule if and only if it has universal domain and is complete, weakly consistent, independent, anonymous, monotonic, and responsive.[11] The important property missing here is consistency. We will investigate if and how the thresholds can be chosen so as to achieve consistency. The properties of independence and monotonicity are equivalent to *strategy-proofness* if each individual i holds *epistemic* preferences, i.e. would like the group to hold beliefs close to A_i, the set of propositions i considers true.[12]

3 A non-classical logic

How should we define the logical interconnections within the language **L** specified in Section 2? Although classical logic gets some entailments right (like $a, a \rightarrow b \vDash b$), its treatment of connection rules is inappropriate, or so I will argue.

Requirements on the representation of connection rules. To reflect the intended meaning of connection rules such as $a \rightarrow b, c \leftrightarrow a, a \rightarrow (b \wedge c)$, the logic should respect the following conditions.

(a) The *acceptance* of a connection rule r establishes exactly the intended logical constraints on atomic propositions, i.e. r is consistent with the "right" sets of atomic and negated atomic propositions. For instance, $a \rightarrow b$ is inconsistent with $\{a, \neg b\}$ but consistent with each of $\{a, b\}$, $\{\neg a, b\}$, $\{\neg a, \neg b\}$.

[11] *Independence*: for all $p \in X$ and all admissible profiles $(A_1, ..., A_n)$, $(A_1^*, ..., A_n^*)$, if $\{i : p \in A_i\} = \{i : p \in A_i^*)$ then $p \in F(A_1, ..., A_n) \Leftrightarrow p \in F(A_1^*, ..., A_n^*)$. *Anonymity*: $F(A_1, ..., A_n) = F(A_{\pi(1)}, ..., A_{\pi(n)})$ for all admissible profiles $(A_1, ..., A_n)$, $(A_{\pi(1)}, ..., A_{\pi(n)})$, where $\pi : N \mapsto N$ is any permutation. *Monotonicity*: for all individuals i and admissible profiles $(A_1, ..., A_n)$, $(A_1, ..., A_i^*, ..., A_n)$ differing only in i's judgment set, if $F(A_1, ..., A_n) = A_i^*$ then $F(A_1, ..., A_i^*, ..., A_n) = A_i^*$. *Responsiveness*: for all $p \in X$ (such that neither p nor $\neg p$ is a tautology) there are admissible profiles $(A_1, ..., A_n)$, $(A_1^*, ..., A_n^*)$ with $p \in F(A_1, ..., A_n)$ and $p \notin F(A_1^*, ..., A_n^*)$. Clearly, quota rules satisfy all seven axioms. Conversely, independence and anonymity imply that the group judgment on any given $p \in X$ depends only on the number $n_p := |\{i : p \in A_i\}|$. This dependence is positive by monotonicity, hence described by an acceptance threshold $m_p \in \{0, 1, ..., n + 1\}$. If p and $\neg p$ are not tautologies, m_p is by responsiveness not 0 and not $n+1$, i.e. $m_p \in \{1, ..., n\}$; and $m_{\neg p} = n - m_p + 1$ by completeness and weak consistency. If p or $\neg p$ is a tautology, we may assume w.l.o.g. that, again, $m_p \in \{1, ..., n\}$ and $m_{\neg p} = n - m_p + 1$.

[12] That is, i weakly prefers the group to hold judgment set A over judgment set B if for all $p \in X$ on which A_i agrees with B, A_i also agrees with A. This condition only partly fixes i's preferences, but it for instance implies that i most prefers i's own judgment set A_i. See Dietrich and List [5].

(b) The *negation* of a (non-degenerate) connection rule r does *not* constrain atomic propositions, i.e. $\neg r$ is consistent with *each* (consistent) set of atomic and negated atomic propositions. For instance, $\neg(a \to b)$ is consistent with each of $\{a, b\}$, $\{a, \neg b\}$, $\{\neg a, b\}$, $\{\neg a, \neg b\}$.

To illustrate (b), consider again the central bank example, where a is "GDP growth will pick up" and b is "inflation will pick up". Consider a board member who believes that $\neg(a \to b)$, i.e. that rising GDP does *not* imply rising inflation. This belief is intuitively perfectly consistent with any beliefs on a and b, i.e. on whether GDP will grow and whether inflation will rise.

The failure of the material implication. Material (bi-)implications (used in classical logic) satisfy (a) but not (b). Consider $a \to b$. Interpreted materially, $a \to b$ is equivalent to $\neg a \vee b$ (not-a or b), and $\neg(a \to b)$ to $a \wedge \neg b$ (a and not-b); so:

- (a) holds because $a \to b$ is inconsistent with $\{a, \neg b\}$ (as desired) and consistent with each of $\{a, b\}, \{\neg a, b\}, \{\neg a, \neg b\}$ (as desired);
- (b) is violated because $\neg(a \to b)$, far from imposing no constraints, is inconsistent with all sets containing $\neg a$ or containing b.

It is well-known that the material interpretation misrepresents the intended meaning of most conditional statements in common language. The (in common language clearly false) statement "if the sun stops shining then we burn" is *true* materially because the sun does *not* stop shining. The material interpretation clashes with intuition because, in common language, "if a then b" is not a statement about the *actual* world, but about whether b holds in hypothetical world(s) where a holds, e.g. worlds where the sun stops shining. "If a then b" thus means "if a *were* true ceteris paribus, then b *would be* true", not "a *is* false or b *is* true".

A conditional logic. A *subjunctive* reading of "$\to$", where the truth value of $a \to b$ depends on b's truth value in possibly non-actual worlds, has been formalised using *possible-worlds semantics*, and more specifically using *conditional logic* which originated from Stalnaker [25] and D. Lewis [13] and is now well-established in non-classical logic. I use a standard version of conditional logic, sometimes denoted C^+ (other versions could also be used). For further reference, e.g. Priest [24].

For comparison, recall that in *classical* logic (not in C^+) $A \subseteq \mathbf{L}$ entails $p \in \mathbf{L}$ if and only if every classical interpretation that makes all $q \in A$ true makes p true, where a *classical interpretation* is simply a ("truth") function $v : \mathbf{L} \to \{T, F\}$ that assigns to each proposition a truth value such that, for all $p, q \in \mathbf{L}$,

- $v(\neg p) = T$ if and only if $v(p) = F$,
- $v(p \wedge q) = T$ if and only if $v(p) = T$ and $v(q) = T$,
- $v(p \to q) = T$ if and only if $v(p) = F$ or $v(q) = T$ (material implication).

This leads to counter-intuitive entailments like $\neg a \vDash a \to b$ and $b \vDash a \to b$, the so-called paradoxes of material implication. In response, the notion of an "interpretation" must be redefined. A C^+*-interpretation* consists of

- a non-empty set W of *(possible) worlds* w;

- for every proposition $p \in \mathbf{L}$ a function $f_p : W \to \mathcal{P}(W)$ ($f_p(w)$ contains the worlds to which "if p were true" refers, i.e. the worlds "similar" to w and with true p);
- for every world $w \in W$ a ("truth") function $v_w : \mathbf{L} \to \{T, F\}$ (that tells what propositions hold in w).

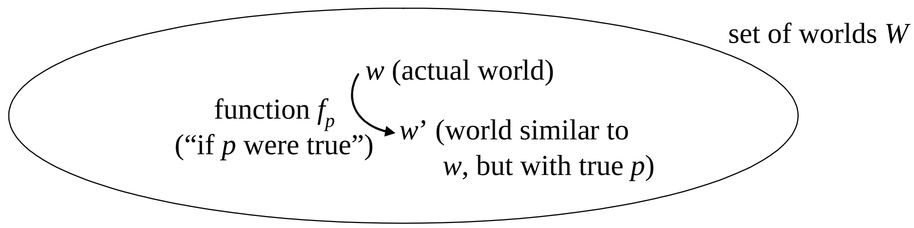

Figure 2: Referring to a non-actual world, in a C^+-interpretation

But not *any* such triple $(W, (f_p), (v_w)) \equiv (W, (f_p)_{p \in \mathbf{L}}, (v_w)_{w \in W})$ may reasonably count as an interpretation: indeed, the meaning of the functions f_p and v_w suggests requiring additional properties. Specifically, such a triple $(W, (f_p), (v_w))$ is defined as a *(C^+-)interpretation* if, for all worlds $w \in W$ and all propositions $p, q \in \mathbf{L}$,

- $v_w(\neg p) = T$ if and only if $v_w(p) = F$ (like in classical logic),
- $v_w(p \wedge q) = T$ if and only if $v_w(p) = T$ and $v_w(q) = T$ (like in classical logic),
- $v_w(p \to q) = T$ if and only if $v_{w'}(q) = T$ for all worlds $w' \in f_p(w)$ (subjunctive implication),
- if $w' \in f_p(w)$ then $v_{w'}(p) = T$ (i.e. p holds in the worlds to which "if p were true" refers),
- if $v_w(p) = T$ then $w \in f_p(w)$ (i.e. if p already holds in w then "if p were true" refers to w).

The truth condition for $p \to q$ (third bullet point) captures the intuitive meaning of implications. "If the sun stops shining then we burn" is false in our world: we do not burn in worlds similar to ours but without the sun shining.

By definition, $A \subseteq \mathbf{L}$ *(C^+-)entails* $p \in \mathbf{L}$ ($A \vDash p$) if, for all interpretations $(W, (f_p), (v_w))$ and all worlds $w \in W$, if all $q \in A$ hold in w then p holds in w (i.e. p holds "whenever" all $q \in A$ hold). For instance, $a, b, (a \wedge b) \to c \vDash c$, but $\neg a \nvDash a \to b$ and $b \nvDash a \to b$ (so C^+ does not suffer the paradoxes of material implication). Recall that $A \subseteq \mathbf{L}$ is consistent if and only if there is no $p \in \mathbf{L}$ with $A \vDash p$ and $A \vDash \neg p$. So

- A is (C^+-)consistent if and only if there is an interpretation $(W, (f_p), (v_w))$ and a world $w \in W$ in which all $q \in A$ hold (i.e. all $q \in A$ "can" hold simultaneously).

So $\{a, \neg a\}$ is inconsistent: if a holds in a word w, $\neg a$ is false in w. And $\{a, \neg(a \to b), b\}$ is consistent (but classically inconsistent): let a and b both hold in w and let $f_a(w)$ contain a world w' in which b is false.[13]

[13] This is an example of why C^+ meets our requirement (b) on the treatment of connection rules. To verify (b) in general, apply Lemma 8 to sets A consisting of negated non-degenerate connection rules and of atomic or negated atomic propositions (and note that (20) does not hold since $A \cap \mathcal{R} = \emptyset$).

4 Simple implication agendas

Given the logic C^+, which quota rules are consistent? I first give an answer for *simple* implication agendas.

Theorem 1 *A quota rule $F_{(m_p)_{p\in X^+}}$ for a simple implication agenda X is consistent if and only if*

$$m_b \leq m_a + m_{a\to b} - n \text{ for all } a \to b \in X. \tag{1}$$

So consistent quota rules do exist: putting $m_p = n$ for all $p \in X^+$ validates (1). But this extreme quota rule is far from the only consistent quota rule; for instance, (1) holds if all atomic propositions $a \in X$ get the same threshold m_a (so all issues are treated symmetrically) and all connection rules $a \to b \in X$ get the unanimity threshold $m_{a\to b} = n$ (so links between issues are very hard to accept).

Some consequences of (1) can be expressed in terms of the network structure of the (simple) implication agenda X (see Figure 1 for examples of network structures). The nodes are the atomic propositions in X, and if $a \to b \in X$ then a is a *parent* of b and b a *child* of a. The notions of *ancestor* and *descendant* follow by transitive closure. By (1), $m_a \geq m_b$ if a is a parent, or more generally an ancestor, of b. In particular, $m_a = m_b$ if a and b are in a cycle, i.e. are ancestors of each other. In short, thresholds of atomic propositions weakly decrease along (descending) paths, and are constant within cycles. Cycles severely restrict the thresholds not only of its member propositions but also of the connection rules $a \to b$ linking them: we must have $m_{a\to b} = n$, as is seen by setting $m_a = m_b$ in (1).

The picture changes radically if we misrepresent the decision problem by using classical logic: then there exists at most one and typically no consistent quota rule $F_{(m_p)_{p\in X^+}}$, where "consistent" now means *classically* consistent and the (universal) domain of $F_{(m_p)_{p\in X^+}}$ now consists of the profiles of complete and *classically* consistent judgment sets.[14] More precisely, the classical counterpart of Theorem 1 is the following result.

Theorem 1* *Defining logical interconnections using classical logic, a quota rule $F_{(m_p)_{p\in X^+}}$ for a simple implication agenda X is consistent if and only if*

$$m_a = n \text{ and } m_{a\to b} = m_b = 1 \text{ for all } a \to b \in X. \tag{2}$$

So there is *no* classically consistent quota rule if X contains a "chain" $a \to b, b \to c$; and there is a *single* (unnatural) one otherwise.

Each of the two theorems can be proven in two steps: step 1 identifies possible types/sources of inconsistency, and step 2 shows that (1) (respectively, (2)) is necessary and sufficient to prevent these types of inconsistency.

More precisely, Theorem 1 follows from the following two lemmas (steps) by noting that any collective judgment set $A \subseteq X$ generated by a quota rule satisfies:

$$A \text{ contains exactly one member of each pair } p, \neg p \in X. \tag{3}$$

[14] Within a simple implication agenda X, the classical logical interconnections are stronger than the C^+ ones: all classically consistent sets $A \subseteq X$ are C^+-consistent but not vice versa (see Lemmas 1, 2). So a consistent quoto rule's (universal) domain and co-domain shrink by moving to classical logic.

Lemma 1 *For a simple implication agenda X, a set $A \subseteq X$ satisfying (3) is consistent if and only if it contains no triple $a, a \to b, \neg b \in X$.*

Lemma 2 *For a simple implication agenda X, a quota rule $F_{(m_p)_{p \in X^+}}$ never accepts any triple $a, a \to b, \neg b \in X$ if and only if (1) holds.*[15]

Analogously, Theorem 1* follows from the following two lemmas (steps).

Lemma 1* *For a simple implication agenda X, a set $A \subseteq X$ satisfying (3) is classically consistent if and only if it contains no triple $a, a \to b, \neg b \in X$ or pair $b, \neg(a \to b) \in X$ or pair $\neg a, \neg(a \to b) \in X$.*

Lemma 2* *For a simple implication agenda with the logical interconnections of classical logic, a quota rule $F_{(m_p)_{p \in X^+}}$ never accepts any triple $a, a \to b, \neg b \in X$ or pair $b, \neg(a \to b) \in X$ or pair $\neg a, \neg(a \to b) \in X$ if and only if (2) holds.*

Lemmas 1 and 1* highlight the difference between non-classical and classical logic: the latter creates two additional types of inconsistency (in simple implication agendas). These additional inconsistencies are artificial; e.g. b is intuitively consistent with $\neg(a \to b)$. By Lemma 2, (1) is necessary and sufficient to exclude all *non-classical* inconsistencies $a, a \to b, \neg b \in X$. But (1) does nothing to prevent the artificial classical inconsistencies.[16] To prevent also these, (1) must be strengthened to (2) by Lemma 2*.

I first prove Lemmas 1 and 1*, in reverse order to start simple.

Proof Lemma 1.* Let X and A be as specified. Clearly, if A contains a triple $a, a \to b, \neg b$ or pair $b, \neg(a \to b)$ or pair $\neg a, \neg(a \to b)$, then A is classically inconsistent. Now suppose A does not contain such a triple or pair. To show A's classical consistency, I define a classical interpretation $v : \mathbf{L} \to \{T, F\}$ that affirms all $p \in A$. Define v by the condition that the only true atomic propositions are those in A. Then all atomic or negated atomic members of A are true. Further, every $a \to b \in A$ is true: as A does not contain the triple $a, a \to b, \neg b$, A either contains b, in which case b is true, hence $a \to b$ is true; or A contains $\neg a$, in which case a is false, hence $a \to b$ is true. Finally, every $\neg(a \to b) \in A$ is true: as A contains neither the pair $b, \neg(a \to b)$ nor the pair $\neg a, \neg(a \to b)$, A contains neither b nor $\neg a$, so that b is false and a true, and hence $a \to b$ is false, i.e. $\neg(a \to b)$ is true.

Proof of Lemma 1. Let X and $A \subseteq X$ be as specified. If $A \subseteq X$ contains a triple $a, a \to b, \neg b$, A is of course (C^+-)inconsistent. Now assume A contains no triple $a, a \to b, \neg b$. To show that A is consistent (though perhaps classically inconsistent), I specify a C^+-interpretation $(W, (f_p), (v_w))$ with a world $\bar{w} \in W$ in which all $p \in A$ hold. Let W contain:

(a) a world $\bar{w}$, in which an atomic proposition a holds if and only if $a \in A$;
(b) for every atomic proposition a, a world w_a ($\neq \bar{w}$) such that

[15] "Never" of course means "for no profile in the (universal) domain of $F_{(m_p)_{p \in X^+}}$".
[16] For instance, the pair $b, \neg(a \to b) \in X$ is collectively accepted if $m_b < m_{a \to b}$ (which (1) allows) and if m_b persons accept the pair $b, a \to b$ and all others accept the pair $\neg b, \neg(a \to b)$.

- $f_a(\bar{w}) = \{w_a\}$ if $a \notin A$ and $f_a(\bar{w}) = \{\bar{w}, w_a\}$ if $a \in A$ (so "if a were true" refers to w_a, and as required by the notion of a C^+-interpretation also to the actual world $\bar{w}$ if a holds there, i.e. if $a \in A$);
- in w_a exactly those atomic propositions b are false for which $\neg(a \to b) \in A$.

We have to convince ourselves that all $p \in A$ hold in $\bar{w}$. All atomic or negated atomic $p \in A$ hold in $\bar{w}$ by (a). Also any negated implication $\neg(a \to b) \in A$ holds in $\bar{w}$: by (b), "if a were true" refers to w_a, in which b is false; whence in $\bar{w}$ $a \to b$ is false, i.e. $\neg(a \to b)$ true. Finally, suppose $a \to b \in A$. I have to show that b holds in all worlds $w \in f_a(\bar{w})$. There are two cases.

Case 1: $a \in A$. Then $f_a(\bar{w}) = \{\bar{w}, w_a\}$ by (b). First, b holds in $\bar{w}$: otherwise $b \notin A$ (by (a)), so that A would contain the triple $a, a \to b, \neg b$, a contradiction. Second, b holds in w_a: otherwise $\neg(a \to b) \in A$ (by (b)), contradicting $a \to b \in A$.

Case 2: $a \notin A$. Then $f_a(\bar{w}) = \{w_a\}$; and b holds in w_a, as just mentioned. ■

I now show Lemmas 2 and 2*, completing the proof of Theorems 1 and 1*.

Proof of Lemma 2. Let $F_{(m_p)_{p \in X^+}}$ be a quota rule for a simple implication agenda X. Take a given triple $a, a \to b, \neg b \in X$. I consider all profiles for which a and $a \to b$ are collectively accepted, and I show that b is collectively accepted (i.e. $\neg b$ rejected) for all such profiles *if and only if* $m_b \leq m_a + m_{a \to b} - n$.

Note first that in all such profiles at least m_a people accept a and at least $m_{a \to b}$ people accept $a \to b$; hence the number of people accepting both these propositions (hence also b) is at least $m_a + m_{a \to b} - n$ (in fact, at least $\max\{m_a + m_{a \to b} - n, 0\}$). Thus, if $m_b \leq m_a + m_{a \to b} - n$, b is in all such profiles accepted by at least m_b people, hence collectively accepted.

For the converse, note that among such profiles there is one such that exactly $\max\{m_a + m_{a \to b} - n, 0\}$ people accept both a and $a \to b$ (hence b) and such that no one else accepts b. If $m_b > m_a + m_{a \to b} - n$, then in this profile less than m_b people accept b, so that b is collectively rejected. ■

Proof of Lemma 2.* Let $F_{(m_p)_{p \in X^+}}$ be a quota rule for a simple implication agenda X, with the *classical* logical interconnections. Lemma 2 (and its proof) also holds under classical logic; so $F_{(m_p)_{p \in X^+}}$ never accepts any triple $a, a \to b, \neg b \in X$ if and only if (1) holds. Further, a given pair $b, \neg(a \to b) \in X$ is never accepted if and only if $m_b \geq m_{a \to b}$: necessity of $m_b \geq m_{a \to b}$ follows from footnote 16, and sufficiency holds because, as b classically entails $a \to b$, $a \to b$ is (in any profile) accepted by at least as many people as b. By an analogous argument, a given pair $\neg a, \neg(a \to b) \in X$ is never accepted if and only if $m_{a \to b} \leq n - m_a + 1$ $(= m_{\neg a})$. In summary, we thus have three inequalities for every $a \to b \in X$: that in (1), $m_b \geq m_{a \to b}$, and $m_{a \to b} \leq n - m_a + 1$. Together these inequalities are equivalent to the condition in (2), as is easily checked. ■

Constructing consistent quota rules. I now discuss how to choose thresholds $(m_p)_{p \in X^+}$ that satisfy (1), for a simple implication agenda X. The notions of a child/parent and a descendant/ancestor are defined above. A *path* is a sequence $(a_1, a_2, ..., a_k)$ in X $(k \geq 2)$ in which each a_j is a parent of a_{j+1} $(j < k)$. X is *acyclic* if it has no cycle, i.e. no path $(a_1, ..., a_k)$ with $a_1 = a_k$. The *depth* of X is $d_X := \sup\{k$

: there is a path in X of length $k\}$, and the *level* of an atomic proposition $a \in X$ is $l_a := \sup\{k :$ there is a path in X of length k ending with $a\}$, interpreted as 1 if no path ends with a. So $a \in X$ has level 1 if it has no parents, level 2 if it has parents all of which have level 1, etc. Figure 3 shows an acyclic simple implication agenda with three levels.

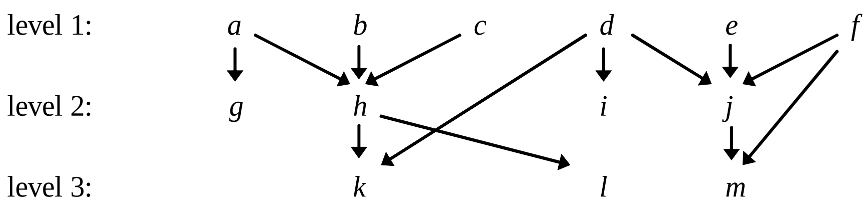

Figure 3: An acyclic simple implication agenda X of depth $d_X = 3$.

How free are we in choosing the thresholds $(m_p)_{p \in X^+}$? Clearly, by (1) the thresholds of atomic propositions must weakly decrease along any path. If X is acyclic and finite (hence of finite depth d_X), $(m_p)_{p \in X^+}$ can be chosen recursively in the following d_X steps.

Step l **(**$= 1, 2, ..., d_X$**):** for all $b \in X$ of level l, choose a threshold $m_b \in \{1, ..., n\}$ and thresholds $m_{a \to b} \in \{1, ..., n\}$ for the parents a of b, such that

$$m_b \leq m_a + m_{a \to b} - n \text{ for all parents } a \text{ of } b. \quad (4)$$

But this procedure may involve choosing many thresholds: in Figure 3, those of 13 atomic propositions and 13 implications! To reduce complexity, one might use

- the same threshold $m = m_{a \to b}$ for all connection rules $a \to b \in X$, where m reflects how easily the group imposes constraints between issues,
- the same threshold m_l for all propositions in X with the same level l ($\in \{1, ..., d_X\}$), where m_l reflects how easily the group accepts level l propositions.

I write such a quota rule as $F_{m,m_1,...,m_{d_K}}$. Here, only $d_X + 1$ parameters must be chosen, e.g., in Figure 3, $3 + 1 = 4$ parameters instead of 26. Applied to quota rules of this type, Theorem 1 yields the following characterisation, by a proof left to the reader.

Corollary 1 *For a finite acyclic simple implication agenda X, a quota rule $F_{m,m_1,...,m_{d_K}}$ is consistent if and only if*

$$m_l \leq m_{l-1} + m - n \textit{ for all levels } l \in \{2, ..., d_X\}. \quad (5)$$

Consistent quota rules of type $F_{m,m_1,...,m_{d_K}}$ can be constructed as follows.

Step 0: choose $m \in \{1, ..., n\}$ such that (i) $m \geq n - (n-1)/(d_X - 1)$.
Step l **(**$= 1, 2, ..., d_X$**):** choose $m_l \in \{1, ..., n\}$ such that (ii) $m_l \geq 1 + (d_X - l)(n - m)$ and (iii) $m_l \leq m_{l-1} + m - n$ if $l > 1$.

The conditions (i)-(iii) follow from Corollary 1: (iii) is obvious, and (i) and (ii) make the choices in future steps possible.[17] For a group of size $n = 10$ and the agenda of Figure 3, a consistent quota rule F_{m,m_1,m_2,m_3} might be chosen as follows.

Step 0: $m = 8$ (note that $8 \geq n - (n-1)/(d_X - 1) = 10 - 9/2 = 5.5$).
Step 1: $m_1 = 8$ (note that $8 \geq 1 + (d_X - 1)(n - m) = 1 + 2 \times 2 = 5$).
Step 2: $m_2 = 6$ (note that $6 \geq 1 + (d_X - 2)(n - m) = 1 + 2 = 3$ and $6 \leq m_1 + m - n = 8 + 8 - 10 = 6$).
Step 3: $m_3 = 4$ (note that $4 \geq 1 + (d_X - 3)(n - m) = 1$ and $4 \leq m_2 + m - n = 6 + 8 - 10 = 4$).

Causal and justificational interpretation. I now offer two interpretations of connection rules, and hence of the kind of decision problems captured by implication agendas. For simplicity, I restrict myself to a *simple* implication agendas X.

First, suppose implications $a \to b \in X$ have a *causal* status: $a \to b$ means that fact a *causes* fact b. So X might contain "*if* the ozone hole has size X *then* global warming will continue" and "*if* global warming will continue *then* species Y will die out". Then X captures a decision problem of forming beliefs about facts and their causal links. A path $(a_1, ..., a_k)$ in X is a *causal chain* (assuming the causal links $a_1 \to a_2, ..., a_{k-1} \to a_k$ hold), and the level of a proposition indicates how "causally fundamental" it is. By an earlier remark, Theorem 1 implies that the acceptance threshold must weakly decrease along any causal chain.

Second, suppose the implications $a \to b \in X$ have a *justificational* (or *evidential* or *indicative*) status: $a \to b$ means that a *indicates* b (a can indicate b without causing b: a wet street indicates rain without causing it). So X captures a decision problem of forming beliefs about claims/statements/hypotheses and their justificational links. Some claims may have a *normative* content, like "a multi-cultural society is desirable" or "option x is better than option y". For instance, an environmental panel might decide on a : "the ozone hole has size larger than X", b : "tax T on kerosine should be introduced", and the justificational link $a \to b$. A path $(a_1, ..., a_k)$ is an "argumentative" chain (assuming the links $a_1 \to a_2, ..., a_{k-1} \to a_k$ hold), and the level of a proposition reflects how "argumentatively fundamental" it is. Often, high level propositions are more concrete and might state that certain collective acts should be taken (a road should be built, a firm downsized, a law amended, etc.), whereas their ancestors describe potential *reasons* or *arguments*, either of a descriptive kind (traffic will increase, demand will fall, etc.) or of a normative kind (multi-culturalism is desirable, etc.). Of course, one may reject a reason $a \in X$, or reject a's status as reason for $b \in X$ (i.e. reject $a \to b$). Again, reasons need at least as high acceptance thresholds as their (argumentative) descendents, e.g. $\frac{3}{4}n$ versus $\frac{1}{2}n$.

5 Other special implication agendas

The difference between using non-classical and using classical logic is now further illustrated by considering two other types of implication agendas X, namely

[17] For instance, without (i) there would be *no* choices of $m_1, ..., m_{d_X}$ satisfying (5).

- *semi-simple* ones: here all connection rules in X are implications $p \to b$ in which b is atomic (such as $(a \wedge c) \to b$ but not $a \to (b \wedge c)$);
- *bi-simple* ones: here all connection rules in X are bi-implications $a \leftrightarrow b$ in which a and b are atomic.

For each of these agenda types, we again perform two steps (analogous to the steps performed for simple implication agendas):

(step 1) we identify possible types/sources of inconsistency in the agenda;
(step 2) we exclude each one by an inequality on thresholds and if possible we simplify the system of inequalities.

This gives a consistency condition for quota rules $F_{(m_p)_{p \in X^+}}$, in analogy to Theorem 1. Again, using instead classical logic leads in step 1 to additional (artificial) types of inconsistency; so that, in analogy to Theorem 1*, at most one (degenerate) quota rule $F_{(m_p)_{p \in X^+}}$ is classically consistent (if X is semi-simple), or even no one (if X is bi-simple). Table 2 summarises the results for each agenda type

Agenda type	A set $A \subseteq X$ satisfying (3) is consistent *iff* it has no subset of type(s)...	$F_{(m_p)_{p \in X^+}}$ is consistent *iff*...
simple	$\{a, a \to b, \neg b\}$ in CL also: $\{b, \neg(a \to b)\}, \{\neg a, \neg(a \to b)\}$	(1) in CL: (2)
semi-simple	$\mathrm{C}(p) \cup \{p \to b, \neg b\}$ in CL also: $\{b, \neg(p \to b)\}, \{\neg a, \neg(p \to b)\}$ $(a \in \mathrm{C}(p))$	(7) in CL: (8)
bi-simple	$\{a, \neg b, a \leftrightarrow b\}, \{\neg a, b, a \leftrightarrow b\}, \{a \leftrightarrow b, \neg(b \leftrightarrow a)\}$ in CL also: $\{a, b, \neg(a \leftrightarrow b)\}, \{\neg a, \neg b, \neg(a \leftrightarrow b)\}$	(9) in CL: never

Table 2: Three types of implication agendas X, their types of inconsistencies, and their consistent quota rules; in non-classical logic and in classical logic ("CL")

In Table 2, the results for simple X were shown in the last section. Regarding semi- or bi-simple X, we have to adapt Lemmas 1 and 2 (or 1* and 2* for classical logic). Let me briefly indicate how this works. First a general remark. The types of inconsistency in step 1 can be (and are in Table 2) identified with certain inconsistent sets $Y \subseteq X$ (which are *minimal inconsistent*, in fact *irreducible*; see Section 7); and the inequality needed in step 2 to exclude Y's acceptance can always be written as

$$\sum_{p \in Y} (n - m_p) < n, \text{ or equivalently } \sum_{p \in Y} m_p > n(|Y| - 1) \tag{6}$$

(see Lemma 5 in Section 7). Intuitively, (6) requires the propositions in Y to have sufficiently high acceptance thresholds to prevent joint acceptance of all $p \in Y$.

First let X be semi-simple. In step 1 we have to consider not only inconsistent sets of type $\{a, a \to b, \neg b\} \subseteq X$ (as for simple X) but also ones like $\{a, c, (a \wedge c) \to b, \neg b\} \subseteq X$. By adapting Lemma 1 to semi-simple agendas, the inconsistent sets in step 1 turn out to be precisely the sets $\mathrm{C}(p) \cup \{p \to b, \neg b\} \subseteq X$. In the proof of Lemma 1, the C^+-interpretation $(W, (f_p), (v_w))$ should be adapted by letting W contain:

(a) a world $\bar{w}$, in which an atomic proposition a holds iff $a \in A$ (as before)

(b) for any conjunction p of atomic propositions a world w_p ($\neq \bar{w}$) such that

- $f_p(\bar{w}) = \{w_p\}$ if $\mathrm{C}(p) \nsubseteq A$ and $f_p(\bar{w}) = \{\bar{w}, w_p\}$ if $\mathrm{C}(p) \subseteq A$ (so "if p were true" refers to world w_p, and also to the actual world $\bar{w}$ if p holds there);
- in w_p exactly those atomic propositions b are false for which $\neg(p \to b) \in A$.

The rest of Lemma 1 – showing that all $p \in A$ hold in $\bar{w}$ – is easily adapted. Using (6), it then follows that a quota rule $F_{(m_p)_{p\in X^+}}$ is consistent if and only if $\sum_{a\in \mathrm{C}(p)\cup\{p\to b,\neg b\}}(n - m_a) < n$ for all $p \to b \in X$; or equivalently

$$\sum_{a\in \mathrm{C}(p)} (n - m_a) + m_b \leq m_{p\to b} \text{ for all } p \to b \in X. \tag{7}$$

Note that this characterisation indeed reduces to Theorem 1 if X is simple.

By contrast, classical logic leads in step 1 to new inconsistent sets of type $\{b, \neg(p \to b)\} \subseteq X$ and $\{\neg a, \neg(p \to b)\}$ with $a \in \mathrm{C}(p)$, as is seen by adapting Lemma 1* (without even having to redefine the classical interpretation $v : \mathbf{L} \to \{T, F\}$). As a result, a quota rule $F_{(m_p)_{p\in X^+}}$ is classically consistent if and only if

$$m_a = n \text{ and } m_{p\to b} = m_b = 1 \text{ for all } p \to b \in X \text{ and all } a \in \mathrm{C}(p). \tag{8}$$

Again, this characterisation reduces to Theorem 1* if X is simple.

Now let X be bi-simple. In step 1, the sets of type $\{a, \neg b, a \leftrightarrow b\}$ or $\{\neg a, b, a \leftrightarrow b\}$ or $\{a \leftrightarrow b, \neg(b \leftrightarrow a\}\}$ capture all types of (non-classical) inconsistency. This can be shown by again adapting Lemma 1 and its proof; when defining the C^+-interpretation $(W, (f_p), (v_w))$, we simply have to replace the second bullet point of (b) by:

- in the world w_a (to which "if a were true" refers), exactly those atomic propositions b are false for which $\neg(a \leftrightarrow b) \in A$ or $\neg(b \leftrightarrow a) \in A$.

By adapting Lemma 2 and its proof (or by using (6) and that $m_{\neg q} = n - m_q + 1$ for all $q \in X$), a quota rule $F_{(m_p)_{p\in X^+}}$ is seen to be consistent if and only if, for all $a \leftrightarrow b \in X$, $m_b \leq m_a + m_{a\leftrightarrow b} - n$ and $m_a \leq m_b + m_{a\leftrightarrow b} - n$ and if also $b \leftrightarrow a \in X$ $m_{a\leftrightarrow b} \geq m_{b\leftrightarrow a}$; which is equivalent to:

$$m_{a\leftrightarrow b} = n \text{ and } m_a = m_b \text{ for all } a \leftrightarrow b \in X. \tag{9}$$

So all bi-implications need the unanimity threshold, and two atomic propositions ("issues") need the same threshold if they are linked by a bi-implication in X or, more generally, by a path of bi-implications in X.

In contrast, classical logic leads (by adapting Lemma 1*) to the additional types of inconsistency $\{a, b, \neg(a \leftrightarrow b)\}, \{\neg a, \neg b, \neg(a \leftrightarrow b)\}$ (which are artificial since negating $a \leftrightarrow b$ shouldn't constrain a's and b's truth values: it shouldn't establish the constraint $a \leftrightarrow \neg b$). This leads (using (6)) to the additional inequalities

$$m_a + m_b + m_{\neg(a\leftrightarrow b)} > 2n \text{ and } m_{\neg a} + m_{\neg b} + m_{\neg(a\leftrightarrow b)} > 2n \text{ for all } a \leftrightarrow b \in X.$$

In this, we have by (9) $m_{\neg(a\leftrightarrow b)} = 1$, so that $m_a + m_b \geq 2n$ and $m_{\neg a} + m_{\neg b} \geq 2n$, hence $m_a = m_b = m_{\neg a} = m_{\neg b} = n$, a contradiction. So there is *no* classically consistent quota rule.

Transforming implication agendas into semi-simple ones. Semi-simple implication agendas are of special interest. While they exclude from the agenda many connection rules – all uni-directional ones with non-atomic consequent and all bi-directional ones – each connection rule of the excluded type can be rewritten, in logically equivalent terms, as a conjunction of connection rules of the non-excluded type $p \rightarrow b$ with atomic b: indeed, each uni-directional connection rule $p \rightarrow q$ is equivalent to the conjunction $\wedge_{b \in \mathrm{C}(q) \backslash \mathrm{C}(p)}(p \rightarrow b)$, and each bi-directional connection rule $p \leftrightarrow q$ is equivalent to the conjunction $\left(\wedge_{b \in \mathrm{C}(q) \backslash \mathrm{C}(p)}(p \rightarrow b)\right) \wedge \left(\wedge_{b \in \mathrm{C}(p) \backslash \mathrm{C}(q)}(q \rightarrow b)\right)$. So every implication agenda X can be transformed into a semi-simple one $\tilde{X}$ by replacing every "non-allowed" connection rule $r \in X^{+}$ by all the "allowed" ones $p \rightarrow b$ of which r is a conjunction (up to logical equivalence). For instance, the implication agenda X given by $X^{+} = \{a, b, c, c \leftrightarrow (a \wedge b)\}$, which models the judges' decision problem in a law suit (see Section 2), can be transformed into the semi-simple implication agenda $\tilde{X}$ given by $\tilde{X}^{+} = \{a, b, c, c \rightarrow a, c \rightarrow b, (a \wedge b) \rightarrow c\}$; under $\tilde{X}$, the judges decide not *en bloc* on $c \leftrightarrow (a \wedge b)$, but separately on whether liability of the defendant implies breach of the contract, whether liability implies validity of the contract, and whether breach of a valid contract implies liability.

Should we conclude from this that all collective decision problems describable by an implication agenda X, like the mentioned one of judges in a law suit, can be remodelled using the corresponding semi-simple implication agenda $\tilde{X}$? And that we could therefore restrict ourselves to the semi-simple case? Not quite, because the change of agenda alters the decision problem. More precisely, it refines (i.e. augments) the decision problem: indeed, from any (complete and consistent) judgment set for $\tilde{X}$ we can always derive a unique one for X, but not vice versa. In the example just given, the judgments on the "new" connection rules $c \rightarrow a, c \rightarrow b, (a \wedge b) \rightarrow c \in \tilde{X}$ together imply a judgment on the "old" one $c \leftrightarrow (a \wedge b) \in X$, but not vice versa because if $c \leftrightarrow (a \wedge b)$ is negated we do not know which one(s) of $c \rightarrow a, c \rightarrow b, (a \wedge b) \rightarrow c$ to negate (we only know that at least one of them must be negated). In summary, it *is* true that the decision problem described by X can be settled by moving to the semi-simple agenda $\tilde{X}$, *but* one thereby settles more and one uses richer in- and output information in the aggregation.

6 General implication agendas

Many implication agendas are of neither of the kinds analysed so far, because they contain connection rules like $a \rightarrow (b \wedge c)$ or $(a \wedge b) \leftrightarrow (a \wedge c)$. Which quota rules are consistent for *general* implication agendas (in the non-classical logic C^{+})? In principle, the above two-step procedure applies again. But, for an agenda class as rich as this one, a so far neglected question becomes pressing: what is it that makes an inconsistent set $Y \subseteq X$ a "type of inconsistency" (in step 1)? Why for instance did we count sets $\{a, a \rightarrow b, \neg b\} \subseteq X$ but not sets $\{a, a \rightarrow b, b \rightarrow c, \neg c\} \subseteq X$ as types of inconsistency for simple implication agendas X? Surely, the set $\mathcal{Y}$ of all types of inconsistency $Y \subseteq X$ must, to enable step 1, be chosen such that

$$\textit{every inconsistent set } A \subseteq X \textit{ satisfying (3) has a subset in } \mathcal{Y}. \tag{10}$$

But usually many choices of $\mathcal{Y}$ satisfy (10). Intuitively, it is useful to choose $\mathcal{Y}$ small and simple. An always possible – but often unduly large – choice of $\mathcal{Y}$ is to include in $\mathcal{Y}$ all *minimal inconsistent* sets $Y \subseteq X$.[18] For simple implication agendas X, we chose $\mathcal{Y} = \{\{a, a \to b, \neg b\} : a \to b \in X\}$, although we could have also included minimal inconsistent sets of type $\{a, a \to b, b \to c, \neg c\} \subseteq X$. We were able to exclude such sets (and still satisfy (10)) because such sets are *reducible* in the following sense. For a set $A \subseteq X$ satisfying (*3*), if $A \supseteq \{a, a \to b, b \to c, \neg c\}$ then, as A contains b or $\neg b$, either $A \supseteq \{b, b \to c, \neg c\}$ or $A \supseteq \{a, a \to b, \neg b\}$, whence A has a subset in $\mathcal{Y}$.

In Section 7, a general method to choose $\mathcal{Y}$ is developed, based on a formalisation of what it means to "reduce" an inconsistent set to a simpler one; $\mathcal{Y}$ then contains *irreducible* sets. Applied to implication agendas, the method yields two kinds of irreducible sets, i.e. two types of inconsistency (as shown in the appendix[19]):

(Ir_+) sets representing an inconsistency between a *non-negated* connection rule and atomic or negated atomic propositions, like $\{a \to (b \wedge c), a, \neg b\}$ or $\{a \leftrightarrow b, \neg a, b\}$;

(Ir_-) sets representing an inconsistency between a *negated* connection rule and non-negated connection rules, like $\{\neg(a \to (b \wedge c)), a \to b, a \to c\}$ or $\{\neg(a \to (b \wedge c \wedge d)), a \to (b \wedge c), a \leftrightarrow d\}$.

In step 2, these irreducible sets yield a system of inequalities whose successive simplification gives the characterisation of Theorem 2 below. This characterisation involves, for every $p \to q \in X$, a particular set $X_{p \to q}$. This set is defined in two steps. First, we form the set

$$X_p := \{s \in \mathbf{L} : p \to s \in X \text{ or } p \leftrightarrow s \in X \text{ or } s \leftrightarrow p \in X\}$$

of all propositions "reachable" from p via (bi-)implications in X. From X_p we then form the set

$$X_{p \to q} := \{S \subseteq X_p : S \text{ is minimal subject to } \mathrm{C}(q) \backslash \mathrm{C}(p) \subseteq \cup_{s \in S} \mathrm{C}(s)\}$$

of all sets $S \subseteq X_p$ that have, and are minimal subject to, this property: each atomic proposition "in" q (but not "in" p) is "in" some $s \in S$. So the sets $S \in X_{p \to q}$ minimally "cover" $\mathrm{C}(q) \backslash \mathrm{C}(p)$.

Evaluating X_p and $X_{p \to q}$ is purely mechanical. As a first example, suppose

$$X^+ = \{a, b, c, a \to b, a \to c, a \to (b \wedge c)\}. \tag{11}$$

Here all three implications have antecedent a, where $X_a = \{b, c, b \wedge c\}$. From X_a we then derive $X_{a \to b}, X_{a \to c}$ and $X_{a \to (b \wedge c)}$. For instance, $X_{a \to b}$ contains $\{b\} \subseteq X_a$ and $\{b \wedge c\} \subseteq X_a$ as both minimally "cover" b, but contains neither $\{c\} \subseteq X_a$ (which fails to "cover" b), nor $\{b, c\} \subseteq X_a$ (which "covers" b *non*-minimally as we can remove c), nor any other set $S \subseteq X_a$. Further, $X_{a \to (b \wedge c)}$ does *not* contain $\{c, b \wedge c\} \subseteq X_a$: although this set "covers" $b \wedge c$ (as all atomic propositions "in" $b \wedge c$ are "in" some $s \in \{c, b \wedge c\}$), it does so non-minimally (as c can be removed); but $X_{a \to (b \wedge c)}$ contains $\{b, c\}$ and $\{b \wedge c\}$ (which "cover" $b \wedge c$ minimally). In summary,

$$X_{a \to b} = \{\{b\}, \{b \wedge c\}\}, X_{a \to c} = \{\{c\}, \{b \wedge c\}\}, X_{a \to (b \wedge c)} = \{\{b, c\}, \{b \wedge c\}\}. \tag{12}$$

[18] Y is *minimal inconsistent* if Y is inconsistent but its proper subsets are consistent.

[19] In fact, each type has two subtypes, one for uni- and one for bi-directional connection rules.

As a second example, suppose

$$X^+ = \{a, b, c, a \to b, a \to (b \wedge c), c \leftrightarrow a\}. \tag{13}$$

The two implications, $a \to b$ and $a \to (b \wedge c)$, both have antecedent a, where $X_a = \{b, c, b \wedge c\}$. From X_a we then derive that:

$$X_{a \to b} = \{\{b\}, \{b \wedge c\}\},\ X_{a \to (b \wedge c)} = \{\{b, c\}, \{b \wedge c\}\}. \tag{14}$$

Sets $X_{p \to q}$ appear in Theorem 2 because they are needed to describe inconsistencies of type (Ir$_-$). Let me give an intuition for why the sets $X_{p \to q}$ relate to inconsistencies of type (Ir$_-$) (details are in the appendix). For the agenda (11), $Y = \{\neg(a \to (b \wedge c)), a \to b, a \to c\}$ is an inconsistency of type (Ir$_-$). Y is inconsistent precisely because the conjuncts of $a \wedge b$ are "covered" by the set of consequents of $a \to b, a \to c \in Y$, i.e. by $\{b, c\}$; in fact, they are so minimally: $\{b, c\} \in X_{a \to (b \wedge c)}$. Another agenda X might have the inconsistency of type (Ir$_-$) $\{\neg(a \to (b \wedge c \wedge d)), a \to (b \wedge c), a \leftrightarrow d\}$. This set is inconsistent precisely because the conjuncts of $b \wedge c \wedge d$ are "covered" by the set of consequents $\{b \wedge c, d\}$; they are so minimally: $\{b \wedge c, d\} \in X_{a \to (b \wedge c \wedge d)}$.

I now state the characterisation result (formally proven in the appendix). As usual, $A \Delta B$ denotes the symmetric difference $(A \backslash B) \cup (B \backslash A)$ of sets A and B.

Theorem 2 *A quota rule $F_{(m_p)_{p \in X^+}}$ for an implication agenda X is consistent if and only if the thresholds satisfy the following:*

(a) *for every $p \to q \in X$,*

$$\sum_{a \in \mathrm{C}(p)} (n - m_a) + \max_{b \in \mathrm{C}(q) \backslash \mathrm{C}(p)} m_b \leq m_{p \to q} \leq n - \max_{S \in X_{p \to q}} \sum_{s \in S : p \to s \in X} (n - m_{p \to s});$$

(b) *for every $p \leftrightarrow q \in X$, (i) $m_{p \leftrightarrow q} = n$, (ii) $m_a = n$ for all $a \in \mathrm{C}(p) \cap \mathrm{C}(q)$, and (iii) m_a is the same for all $a \in \mathrm{C}(p) \Delta \mathrm{C}(q)$ and equals n if $|\mathrm{C}(p) \Delta \mathrm{C}(q)| \geq 3$.*

Theorem 2 characterises consistent quota rules by complicated (in)equalities. A rough interpretation is:

- inconsistencies of type (Ir$_+$) are prevented by the LHS inequalities of (a) and by (b);
- given the LHS inequalities of (a) and (b), inconsistencies of type (Ir$_-$) are prevented by the RHS inequalities in (a).

More detailed clues to understand the conditions (a) and (b) are given at the section end, drawing on the insights gained above on the simple, semi-simple and bi-simple case.

In practice, the system (a)&(b) often simplifies. Part (a) or part (b) drops out if X contains no uni- or no bi-directional connection rules, respectively. If X is simple, semi-simple or bi-simple, (a)&(b) reduces to the conditions derived earlier (namely (1), (7) or (9), respectively).[20] Further, the system (a)&(b) may simplify once the

[20] If X is simple, this is so because (b) drops out and because in (a) the RHS inequality holds trivially (by $X_{p \to q} = \{\{q\}\}$) and the LHS inequality reduces to $n - m_p + m_q \leq m_{p \to q}$.

concrete sets $X_{p\to q}$, $p \to q \in X$, are inserted, possibly resulting in a simpler set of conditions that offers an intuition for the size and structure of the space of possible threshold assignments. The next example demonstrate this.

Example. Consider the agenda in (13). Which thresholds $(m_p)_{p\in X^+}$ guarantee consistency? By Theorem 2, three conditions must hold: one for $a \to b$ (part (a)), one for $a \to (b \wedge c)$ (part (a)), and one for $c \leftrightarrow a$ (part (b)). The three conditions are:

$$\begin{cases} n - m_a + m_b \le m_{a\to b} \le n - \max_{S\in X_{a\to b}} \sum_{s\in S: a\to s\in X} (n - m_{a\to s}) \\ n - m_a + \max\{m_b, m_c\} \le m_{a\to(b\wedge c)} \le n - \max_{S\in X_{a\to(b\wedge c)}} \sum_{s\in S: a\to s\in X} (n - m_{a\to s}) \\ m_{c\leftrightarrow a} = n \text{ and } m_a = m_c. \end{cases} \tag{15}$$

In this system, the upper bounds of $m_{a\to b}$ and $m_{a\to(b\wedge c)}$ – I call them $B_{a\to b}$ and $B_{a\to(b\wedge c)}$, respectively – should be computed by inserting the sets $X_{a\to b}$ and $X_{a\to(b\wedge c)}$ as given in (14). Then $B_{a\to b}$ and $B_{a\to(b\wedge c)}$ greatly simplify (and turn out to be equal) because each summation "$\sum_{s\in S: a\to s\in X}$" runs over just one term:

$$\begin{aligned} B_{a\to b} &= n - \max\left\{ \sum_{s\in\{b\}: a\to s\in X} (n - m_{a\to s}), \sum_{s\in\{b\wedge c\}: a\to s\in X} (n - m_{a\to s}) \right\} \\ &= n - \max\{n - m_{a\to b}, n - m_{a\to(b\wedge c)}\} = \min\{m_{a\to b}, m_{a\to(b\wedge c)}\}, \\ B_{a\to(b\wedge c)} &= n - \max\left\{ \sum_{s\in\{b,c\}: a\to s\in X} (n - m_{a\to s}), \sum_{s\in\{b\wedge c\}: a\to s\in X} (n - m_{a\to s}) \right\} \\ &= n - \max\{n - m_{a\to b}, n - m_{a\to(b\wedge c)}\} = \min\{m_{a\to b}, m_{a\to(b\wedge c)}\}. \end{aligned}$$

So, in the system (15), the RHS inequalities on the first two lines are jointly equivalent to $m_{a\to b} = m_{a\to(b\wedge c)}$. By $m_a = m_c$, the LHS inequality on the second line implies $n - m_a + \max\{m_b, m_a\} \le m_{a\to(b\wedge c)}$, and so $\max\{n - m_a + m_b, n\} \le m_{a\to(b\wedge c)}$; which (by $m_{a\to(b\wedge c)} \le n$) implies that $m_{a\to(b\wedge c)} = n$, and that $n - m_a + m_b \le n$, i.e. $m_b \le m_a$. Using all this, the system (15) is equivalent to:

$$m_b \le m_a = m_c \text{ and } m_{a\to b} = m_{a\to(b\wedge c)} = m_{c\leftrightarrow a} = n.$$

This is an example of how the presence of a *bi*-directional connection rule r in X can drastically narrow down the possibility space, especially relative to thresholds of r, of atomic propositions "in" r, and of connection rules logically related to r.

I now record two corollaries of Theorem 2. First, a possibility result follows.[21]

Corollary 2 *For an implication agenda X, there exists*

(i) *a consistent quota rule $F_{(m_p)_{p\in X^+}}$ (hence a consistent, complete, independent, anonymous, monotonic and responsive aggregation rule with universal domain);*

[21] See Section 2 and footnote 11 for the conditions listed in part (i). By a different proof, part (i) holds more generally for any agenda X for which each $p \in X^+$ is atomic or a connection rule (where, unlike for implication agendas, the atomic propositions in X^+ may differ from those contained in the connection rules in X^+).

(ii) *a single consistent quota rule* $F_{(m_p)_{p \in X^+}}$ *with identical thresholds* m_p, $p \in X^+$, *namely the quota rule with a unanimity threshold* $m_p = n$ *for all* $p \in X^+$.

Proof. As (ii) implies (i), I only show (ii). Let X be an implication agenda and $F_{(m_p)_{p \in X^+}}$ a quota rule with identical thresholds $m_p = m$ ($\in \{1, ..., n\}$). If $m = n$ then (a)&(b) hold, implying consistency. Conversely, assume consistency. So (a)&(b) hold. X contains a $p \to q$ or a $p \leftrightarrow q$ (otherwise X would be empty, hence not an agenda). In the second case, $m = n$ by (b). In the first case, the LHS inequality in (a) implies $\sum_{a \in \mathrm{C}(p)} (n - m) + m \leq m$, whence again $m = n$. ■

So there is possibility – but how large is it? That is, how much freedom does Theorem 2 leave us in the choice of thresholds? As I now show, *paths* and *cycles* in X impose rather severe restrictions. Extending earlier definitions from simple to general implication agendas, consider the network over the atomic propositions in X, where an atomic proposition $a \in X$ a *parent* of another one $b \in X$ if there is a $p \to q \in X$ or a $p \leftrightarrow q \in X$ or a $q \leftrightarrow p \in X$ such that $a \in \mathrm{C}(p)$ and $b \in \mathrm{C}(q) \backslash \mathrm{C}(p)$. Parenthood yields the notion of an *ancestor* by transitive closure. A *path* is a sequence $(a_1, ..., a_k)$ $(k \geq 2)$ where a_j is a parent of a_{j+1} for all $j < k$; it is a *cycle* if $a_1 = a_k$.

Corollary 3 *Let* $F_{(m_p)_{p \in X^+}}$ *be a consistent quota rule for an implication agenda* X.

(i) *If* $a \in X$ *is an ancestor of* $b \in X$ *then* $m_a \geq m_b$.

(ii) *If* $a, b \in X$ *occur in a cycle (i.e. are ancestors of each other) then* $m_a = m_b$ *and* $m_{p \to q} = n$ *for all* $p \to q \in X$ *with* $a \in \mathrm{C}(p)$ *and* $b \in \mathrm{C}(q) \backslash \mathrm{C}(p)$.

Proof. Let X and $F_{(m_p)_{p \in X^+}}$ be as specified.

(i) Let $a \in X$ be a parent of $b \in X$ (obviously it suffices to consider this case). Then $a \in \mathrm{C}(p)$ and $b \in \mathrm{C}(q) \backslash \mathrm{C}(p)$, where $p \to q \in X$ or $p \leftrightarrow q \in X$ or $q \leftrightarrow p \in X$. In the last two cases, (b) implies $m_a \geq m_b$. In the first case, the LHS inequality in (a) implies $(n - m_a) + m_b \leq m_{p \to q}$, so $m_b \leq m_{p \to q} - n + m_a \leq m_a$.

(ii) Let a, b be as specified. By (i) $m_a \leq m_b$ and $m_b \leq m_a$, hence $m_a = m_b$. Now let $p \to q$ be as specified. By the LHS inequality in (a), $(n - m_a) + m_b \leq m_{p \to q}$, hence (by $m_a = m_b$) $m_{p \to q} = n$. ■

An intuition for Theorem 2. Our earlier insights about simple, semi-simple and bi-simple implication agendas offer some clues to understand Theorem 2, more precisely to understand the necessity of (b) and of the LHS of (a). General implication agendas X go in three ways beyond simple ones: (i) implications $p \to q \in X$ may have non-atomic antecedent p; (ii) implications $p \to q \in X$ may have non-atomic consequent q; (iii) X may contain bi-implications $p \leftrightarrow q$.

Here, (i) reminds of semi-simple agendas. And indeed, the LHS of (a), for which (i) is responsible, is closely related to our earlier characterisation (7) of consistent quota rules for semi-simple agendas. To see why, suppose first that in (a) $p \to q$ has atomic consequent q. Then the LHS of (a) coincides with the inequality in (7). Now suppose q is non-atomic. Then $p \to q$ is logically equivalent to the conjunction $\wedge_{b \in \mathrm{C}(q) \backslash \mathrm{C}(p)} p \to b$, and the LHS of (a) is equivalent to applying the inequality in (7) to all implications $p \to b, b \in \mathrm{C}(q) \backslash \mathrm{C}(p)$.

Further, (iii) reminds of bi-simple agendas. Part (b), for which (iii) is responsible, is indeed closely related to our characterisation (9) of consistent quota rules for bi-simple agendas. If in (b) both p and q are atomic, (b) is equivalent to the condition in (9). If p and/or q is non-atomic, (b) is the right generalisation of (9), as formally shown in the appendix.

Finally, (ii) is the aspect in which general implication agendas go substantially beyond both semi- and bi-simple ones. It is responsible for the (complex) RHS inequalities in (a). These inequalities are needed because (ii) introduces new types of inconsistency like $\{a \to b, a \to c, \neg(a \to (b \wedge c))\}$.

7 An abstract characterisation result

A central issue so far was that each agenda X has its own types/sources of inconsistency $Y \subseteq X$ (e.g. the sets $\{a, a \to b, \neg b\} \subseteq X$ if X is a simple implication agenda). What exactly are "types/sources" of inconsistency? They are *irreducible* sets $Y \subseteq X$, as made precise now. I introduce an abstract *simplicity* relation between inconsistent sets $Y \subseteq X$, which allows one to simplify inconsistent sets, which yields *irreducible* sets. I do this in full generality, i.e. independently of implication agendas and the particular logic C^+. This gives rise to an abstract characterisation result of which all above characterisations are applications. The notion of irreducible sets generalises a special irreducibility notion introduced by Dietrich and List [8]; it also generalises minimal inconsistent sets (which are based on set-inclusion rather than a general simplicity relation), and for this reason the abstract characterisation result generalises the characterisation by the "intersection property" in Nehring and Puppe [18, 19].

To avoid unnecessary restrictions to special judgment aggregation problems, we adopt Dietrich's [3] general logics framework in this section: let $X \subseteq \mathbf{L}$ be an arbitrary agenda of propositions from any formal language $\mathbf{L}$ with well-behaved logical interconnections.[22] Further, let $\mathcal{I}$ be the set of all inconsistent sets $Y \subseteq X$.

Given the intended purpose, I will define irreducibility in such a way that

$$\textit{every inconsistent and complete set } A \subseteq X \textit{ has an irreducible subset.} \tag{16}$$

This property ensures that collective consistency holds if and only if no irreducible set is ever collectively accepted. Property (16) is the analogue of the property (10) underlying the 2-step procedure in earlier sections. Of course, we could achieve (16) by simply defining "irreducible" as "minimal inconsistent", since any inconsistent set $A \subseteq X$ has a minimal inconsistent subset. But this would often create a large number of irreducible sets (hence many redundant inequalities in step 2). The irreducibility notion I introduce depends on a parameter: the *simplicity* notion used. Under a certain (extreme) simplicity notion, "irreducible" will coincide with "minimal inconsistent"; other simplicity notions lead to fewer irreducible sets.

I now define simplicity (from which I later define irreducibility). Suppose we have a notion of simplicity of sets in $\mathcal{I}$ given by a binary relation $<$ on $\mathcal{I}$, where "$Z < Y$"

[22] The well-behavedness can be expressed either in terms of the entailment notion $\vDash$ (conditions L1-L3 in Dietrich [3]) or in terms of the inconsistency notion (conditions I1-I3 in Dietrich [3]) (assuming that both notions are interdefinable; see Section 2). Stated in terms of the consistency notion, the three conditions are: (I1) sets $\{p, \neg p\} \subseteq \mathbf{L}$ are inconsistent; (I2) subsets of consistent sets are consistent; (I3) the empty set $\emptyset$ is consistent, and each consistent set $A \subseteq \mathbf{L}$ has a consistent superset $B \subseteq \mathbf{L}$ containing a member of each pair $p, \neg p \in \mathbf{L}$. If the agenda X is infinite, I also assume the logic to be compact: every inconsistent set $A \subseteq \mathbf{L}$ has a finite inconsistent subset. All this holds for C^+ and most familiar logics, including propositional and predicate logics, classical and non-classical logics, with the important exception of non-monotonic logics.

is interpreted as "Z is simper than Y". There is much freedom in how to specify $<$ (the goal being to obtain "nice" irreducible sets, as explained later). For instance, we might define $<$ by $Z < Y :\Leftrightarrow |Z| < |Y|$ (i.e. "simpler" means "smaller"), or by $Z < Y :\Leftrightarrow Z \subsetneq_{\text{finite}} Y$ (i.e. "simpler" means to be a proper finite subset). I place only two restrictions on the simplicity notion:

Proper subsets are simpler: for all $Y, Z \in \mathcal{I}$, if $Z \subsetneq_{\text{finite}} Y$ then $Z < Y$.[23]

No infinite simplification chains: $<$ is *well-founded*, i.e. there is no infinite sequence $(Y_k)_{k=1,2,...}$ in $\mathcal{I}$ such that $Y_{k+1} < Y_k$ for all $k = 1, 2, ...$[24]

A *simplicity relation* is a binary relation $<$ on $\mathcal{I}$ with these properties. For instance the two relations $<$ just mentioned are simplicity relations.

Suppose we have chosen a simplicity relation $<$. Then (16) holds for the following reason. Starting from an arbitrary complete set $A \in \mathcal{I}$, one can find a finite sequence of simplifications $A = Y_1 > Y_2 > ...$ such that the simplified sets $Y_1, Y_2, ...$ all remain subsets of A and the sequence terminates with an irreducible set (as defined below). An example is helpful. Suppose $A = Y_1$ is infinite. Then in a first simplification step we can move to a finite inconsistent subset $Y_2 \subseteq A$ (which exists since the logic is compact or X is finite; see footnote 22). To bring the example into familiar terrain, assume that the agenda is an implication agenda and that $Y_2 = \{a, a \rightarrow (b_1 \wedge ... \wedge b_5), b_6, (b_1 \wedge .... \wedge b_6) \leftrightarrow c, \neg c\}$. In the next simplification step, there are two cases.

Case 1: If all of $b_1, ..., b_5$ are in A, the inconsistent set $Y_3 := \{b_1, ..., b_6, (b_1 \wedge .... \wedge b_6) \leftrightarrow c, \neg c\}$ is a subset of A.

Case 2: If not all of $b_1, ..., b_5$ are in A, say $b_j \notin A$, then $\neg b_j \in A$ (as A is complete), and so the inconsistent set $Y_3' := \{a, a \rightarrow (b_1 \wedge ... \wedge b_5), \neg b_j\}$ is a subset of A.

Of course, the new subset of A (Y_3 or Y_3') is not under all simplicity notions $<$ simpler than Y_2: for instance, we have $Y_3 \nless Y_2$ if $<$ is the "smaller than" relation (i.e. $Z < Y \Leftrightarrow |Z| < |Y|$). There is however an obvious simplicity notion for which both Y_3 and Y_3' *are* simpler than Y_2: they contain fewer connection rules than Y_2 (namely one instead of two). Indeed the application to implication agendas (in the appendix) will use a simplicity relation $<$ that lexicographically prioritises minimising the number of (possibly negated) connection rules over minimising the number of (possibly negated) atomic propositions, thereby ensuring that $Y_3 < Y_2$ and $Y_3' < Y_2$.

The set Y_3 is obtained from Y_2 in a particular manner: I have taken in "new" propositions (namely $b_1, ..., b_5$) each of which is logically entailed by some set of "old" propositions, namely by $V = \{a, a \rightarrow (b_1 \wedge ... \wedge b_5)\} \subseteq Y_2$. The fact that each "new" proposition b_j is entailed by a $V \subseteq Y_2$ has the important consequence that a simplification of Y_2 into a subset of A is possible whether or not A contains all "new" propositions: if A does then $Y_3 \subseteq A$, and if A does not contain the "new" proposition

[23] $Z \subsetneq_{\text{finite}} Y$ stands for $Z \subsetneq Y \& |Z| < \infty$. "$\subsetneq_{\text{finite}}$" can be replaced throughout by "$\subsetneq$" if one assumes a finite agenda.

[24] $<$ need not be connected, nor even transitive (if $<$ also satisfies these conditions, $<$ is a well-order). Note that well-foundedness implies asymmetry (i.e. if $Z < Y$ then $Y \nless Z$), hence irreflexivity. Further, given asymmetry and transitivity, $<$ is well-founded if and only if every set $\emptyset \neq \mathcal{J} \subseteq \mathcal{I}$ on which $<$ is connected has a least element (i.e. a $Z \in \mathcal{J}$ with $Z < Y$ for all $Y \in \mathcal{J} \backslash \{Z\}$).

b_j then $V \cup \{\neg b_j\} = Y_3' \subseteq A$. In the latter case, what is it that allows us to simplify Y_2 into $V \cup \{\neg b_j\}$? First, the entailment $V \vDash b_j$ guarantees us that $V \cup \{\neg b_j\}$ is indeed an inconsistent set, i.e. is in the range $\mathcal{I}$ of the simplicity relation. But this alone does not suffice: $V \cup \{\neg b_j\}$ must actually be simpler than Y_2. In summary, the following properties of the set Y_3 ensure that Y_2 can be simplified (into Y_3 or another set): Y_3 is simpler than Y_2, and moreover each "new" proposition $p \in Y_3 \backslash Y_2$ is entailed by a set of "old" propositions $V \subseteq Y_2$ such that $V \cup \{\neg p\}$ is simpler than Y_2. In this case I call Y_3 a *reduction* of Y_2, as formally defined now.[25]

Definition 1 *Given a simplicity relation* $<$*,*

(i) $Z \in \mathcal{I}$ *is a* ($<$-)reduction of $Y \in \mathcal{I}$ *(and* Y *is* ($<$-)reducible *to* Z*) if* $Z < Y$ *and moreover each* $p \in Z \backslash Y$ *is entailed by some* $V \subseteq Y$ *satisfying* $V \cup \{\neg p\} < Y$*;*

(ii) $Y \in \mathcal{I}$ *is* ($<$-)irreducible *if it has no reduction; let* $\mathcal{IR}_< := \{Y \in \mathcal{I} : Y$ *is* $<$*-irreducible*$\}$ *(the set of of minimal elements of the reduction relation).*

The art is to use a simplicity relation $<$ that allows sufficiently many (and the "right") simplifications so as to give few and elegant irreducible sets (hence a simple characterisation of collective consistency). Let me take up the two example above.

Example 1 (being simpler as being a subset). Let $<$ be $\subsetneq_{\text{finite}}$. Then reduction coincides with simplification: $Z \in \mathcal{I}$ is a reduction of $Y \in \mathcal{I}$ if and only if $Z \subsetneq_{\text{finite}} Y$. So Y is irreducible if and only if Y is a *minimal inconsistent* set, i.e. $\mathcal{IR}_< = \mathcal{MI}$ where

$$\mathcal{MI} := \{Y \in \mathcal{I} : \text{no proper subset of } Y \text{ is in } \mathcal{I}\}.$$

It can be shown that if X is a *simple* implication agenda then the set $\mathcal{IR}_< = \mathcal{MI}$ consists of all sets $Y \subseteq X$ of type $Y = \{p, \neg p\}$ or type

$$Y = \{a_1, a_1 \to a_2, ..., a_{k-1} \to a_k, \neg a_k\} \ (a_1, ..., a_k \text{ pairwise distinct, } k \geq 2). \quad (17)$$

Example 2 (being simpler as being smaller). Let $Z < Y :\Leftrightarrow |Z| < |Y|$. Then reduction is equivalent to Dietrich and List's [8] special reduction notion (see footnote 25). If X is again a simple implication agenda, $\mathcal{IR}_<$ is now much smaller than in Example 1: $\mathcal{IR}_<$ can be shown to consist of all sets $Y \subseteq X$ of type $\{p, \neg p\}$, or of type (17) with $k = 2$ (i.e. of type $\{a, a \to b, \neg b\}$, like in Lemma 1). To see why sets of type (17) are *not* irreducible if $k > 2$, note that such a set Y is reducible for instance to $Z := \{a_{k-1}, a_{k-1} \to a_k, \neg a_k\}$, because $|Z| < |Y|$ and a_{k-1} is entailed by $V := \{a_1, a_1 \to a_2, ..., a_{k-2} \to a_{k-1}\}$ where $|V \cup \{\neg a_{k-1}\}| < |Y|$. As a different agenda, consider a standard strict preference aggregation problem with a set of options $K \neq \emptyset$. This can be represented by the agenda $X_K := \{xPy, \neg xPy : x, y \in K\}$ in a suitable predicate logic with a binary *predicate* P for strict preference, a set of *constants* K for options, and a set of *axioms* containing the rationality conditions on strict linear orders, including for instance the transitivity axiom $(\forall v_1)(\forall v_2)(\forall v_3)((v_1 P v_2 \wedge v_2 P v_3) \to v_1 P v_3)$ (see Dietrich and List [7]; also List and Pettit [16]). Dietrich and List [8] call a set $Y \subseteq X_K$ a "k-cycle" ($k \geq 1$) if it has the form

$$Y = \{x_1 P x_2, x_2 P x_3, ..., x_{k-1} P x_k, x_k P x_1\} \ (x_1, ..., x_k \in K \text{ pairwise distinct}), \quad (18)$$

[25] In the special case that $<$ is defined by $Y < Z :\Leftrightarrow |Y| < |Z|$, this definition of reduction (and irreducibility) becomes equivalent to that introduced for different purposes by Dietrich and List [8]. Proposition 1 generalises one of their results. The present notion of reduction is more flexible and general, as sets may be simplified in other ways than through decreasing their size.

or arises from such a set by replacing one or more of the members xPy by the logically equivalent proposition $\neg yPx$. They show that the irreducible sets are the k-cycles with $k \leq 3$. To see why for $k > 3$ a k-cycle is not irreducible (though minimal inconsistent), note that a set Y of type (18) with $k \geq 4$ is reducible to $Z := \{x_1Px_2, x_2Px_3, x_3Px_1\}$ (a 3-cycle), because $|Z| < |Y|$ and x_3Px_1 is entailed by $V := \{x_3Px_4, x_4Px_5, ..., x_kPx_1\}$ where $|V \cup \{\neg x_3Px_1\}| < |Y|$.

To understand the properties of reduction better, let me record two lemmas.

Lemma 3 *Given any simplicity relation $<$, the reduction relation is itself a simplicity relation, that is:*

(i) *for all $Y, Z \in \mathcal{I}$, if $Z \subsetneq_{\text{finite}} Y$ then Z is a reduction of Y;*

(ii) *there is no infinite sequence $(Y_k)_{k=1,2,...}$ in $\mathcal{I}$ such that Y_{k+1} is a reduction of Y_k for all $k = 1, 2, ...$*

Proof. Both parts follow immediately from the analogous properties of $<$. ■

Lemma 4 (i) *For any simplicity relation $<$, $\mathcal{IR}_< \subseteq \mathcal{MI}$, and if $< = \subsetneq_{\text{finite}}$ then $\mathcal{IR}_< = \mathcal{MI}$.*

(ii) *For any simplicity relations $<$ and $<'$, if $<$ is a subrelation of $<'$ then $\mathcal{IR}_{<'} \subseteq \mathcal{IR}_<$.*

Proof. (i) Let $<$ be a simplicity relation. For all $Y \in \mathcal{I}$, if $Y \notin \mathcal{MI}_<$ then Y has an inconsistent proper subset Z, which we can choose finite by compactness of the logic. By Lemma 3 Y is reducible to Z, whence $Y \notin \mathcal{IR}_<$.

(ii) If $<$ and $<'$ are simplicity relations and $<$ is a subrelation of $<'$, then $<$-reduction is a subrelation of $<'$-reduction, and so $\mathcal{IR}_{<'} \subseteq \mathcal{IR}_<$. ■

Lemma 4 gives a general idea on how the set of irreducible sets $\mathcal{IR}_<$ depends on the simplicity notion $<$ used. The finer $<$ is, i.e. the more simplifications are allowed, the more reductions are allowed, and so the smaller $\mathcal{IR}_<$ is (see part (ii)). The coarsest choice of $<$ is $\subsetneq_{\text{finite}}$; then the only reductions are those to finite proper subsets, and $\mathcal{IR}_<$ is maximal: $\mathcal{IR}_< = \mathcal{MI}$, whereas in general $\mathcal{IR}_< \subseteq \mathcal{MI}$ (see part (i)).

I now prove the central property (16) announced earlier: every inconsistent and complete judgment set $A \subseteq X$ has an irreducible subset (reachable from A via finitely many simplifications).[26]

Proposition 1 *Given any simplicity relation $<$, every inconsistent and complete set $A \subseteq X$ has a subset in $\mathcal{IR}_<$.*

So, by Example 2 above, if X is a simple implication agenda then any inconsistent and complete set $A \subseteq X$ has a subset of type $\{p, \neg p\}$ or $\{a, a \to b, \neg b\}$ (as also shown in Lemma 1); and if X is instead the preference agenda X_K, A has a subset that is a k-cycle with $k \leq 3$ – a well-known result of social choice theory since A corresponds to a connected strict preference relation $\succ$ on K with rationality violation.

[26] The condition "proper subsets are simpler" on the simplicity relation $<$ may be dropped in Proposition 1 but not in Theorem 3.

Proof. Let $<$ and A be as specified. Assume for a contradiction that A has no subset in $\mathcal{IR}_<$. I recursively define a sequence $(Y_k)_{k=1,2,...}$ of inconsistent subsets of A such that $Y_{k+1} < Y_k$ for all k. This contradicts the well-foundedness of $<$.

First, put $Y_1 := A$, which is indeed an inconsistent subset of A.

Second, suppose Y_k is already defined. By assumption, Y_k is reducible, say to $Z \in \mathcal{I}$. First assume $Z \subseteq Y_k$. Letting $Y_{k+1} := Z$, it is true that Y_{k+1} is a subset of A (as $Y_{k+1} \subseteq Y_k \subseteq A$) and that $Y_{k+1} < Y_k$ (as Y_{k+1} is a reduction of Y_k). Now suppose $Z \nsubseteq A$. Then there is a $p \in Z \backslash A$. As Z is a reduction of Y_k, there is a $V \subseteq Y_k$ that entails p with $V \cup \{\neg p\} < Y_k$. Letting $Y_{k+1} := V \cup \{\neg p\}$, we have $Y_{k+1} < Y_k$, and $Y_{k+1} \subseteq A$ because $V \subseteq Y_k \subseteq A$ and because $\neg p \in A$ since $p \notin A$ and A is complete. ■

By Proposition 1, a quota rule $F_{(m_p)_{p \in X^+}}$ is consistent if and only if it never accepts any $Y \in \mathcal{IR}_<$. The following lemma tells which inequality we must impose to achieve this.

Lemma 5 *For every minimal inconsistent (hence every irreducible) set $Y \subseteq X$, a quota rule $F_{(m_p)_{p \in X^+}}$ never accepts all $p \in Y$ if and only if $\sum_{p \in Y}(n - m_p) < n$ (where $m_{\neg p} := n - m_p + 1$ for all $p \in X^+$).*

Proof. Consider a minimal inconsistent $Y \subseteq X$ and a quota rule $F := F_{(m_p)_{p \in X^+}}$.

First assume $\sum_{p \in Y}(n - m_p) \geq n$. Then N can be partitioned into (possibly empty) subgroups $N^p, p \in Y$, of size $|N^p| \leq n - m_p$. Construct a profile $(A_1, ..., A_n)$ of complete and consistent judgment sets such that, for all $p \in Y$, the people in N^p reject just p out of Y, i.e. $A_i \supseteq Y \backslash \{p\}$ for all $i \in N^p$; such A_i's exist as $Y \backslash \{p\}$ is consistent. Then $Y \subseteq F(A_1, ..., A_n)$ (as desired) since the number of people accepting a $p \in Y$ is $n - |N^p| \geq n - (n - m_p) = m_p$.

Conversely, suppose that F has an outcome $F(A_1, ..., A_n) \supseteq Y$. I show that $\sum_{p \in Y}(n - m_p) \geq n$. For all $p \in Y$, put $n_p := |\{i : p \in A_i\}|$; hence $|\{i : p \notin A_i\}| = n - n_p$. So

$$|\{(p, i) \in Y \times N : p \notin A_i\}| = \sum_{p \in Y}(n - n_p).$$

As no A_i contains all $p \in Y$, $|\{(p, i) \in Y \times N : p \notin A_i\}| \geq n$, i.e. $\sum_{p \in Y}(n - n_p) \geq n$. So, as for all $p \in Y$ we have $n_p \geq m_p$ (by $p \in F(A_1, ..., A_n)$), $\sum_{p \in Y}(n - m_p) \geq n$. ■

Proposition 1 and Lemma 5 imply the desired characterisation result.

Theorem 3 *For any simplicity relation $<$, a quota rule $F_{(m_p)_{p \in X^+}}$ is consistent if and only if*

$$\sum_{p \in Y}(n - m_p) < n \text{ for all } Y \in \mathcal{IR}_< \text{ (where } m_{\neg p} := n - m_p + 1 \ \forall p \in X^+).$$

Theorem 3 generalises the anonymous case of the "intersection property" result in Nehring and Puppe [18, 19]. This result makes no reference to a simplicity relation and uses $\mathcal{MI}$ instead of $\mathcal{IR}_<$. Hence it follows from Theorem 3 by choosing $<$ such that $\mathcal{IR}_< = \mathcal{MI}$, i.e. by choosing $<$ as the coarsest simplicity relation $\subsetneq_{\text{finite}}$. A non-anonymous variant of Theorem 3 can be derived similarly, generalising the non-anonymous intersection property result.[27] One might wonder whether one could also generalise the (anonymous or non-anonymous) intersection property result in Dietrich and List [6] which requires no collective completeness[28], again by using irreducible sets instead of minimal inconsistent sets. No straightforward generalisation works, since the completeness assumption is essential in Proposition 1.

In general, the finer the simplicity relation $<$ is chosen, the smaller $\mathcal{IR}_<$ becomes, and hence the "slimmer" Theorem 3's characterisation becomes since redundant inequalities are avoided. The question of how much smaller than $\mathcal{MI}$ the set $\mathcal{IR}_<$ can get (and hence how much "slimmer" than the intersection property result Theorem 3's characterisation can get) depends on the concrete agenda X. In Example 2 above, $\mathcal{IR}_<$ gets significantly smaller than $\mathcal{MI}$. Note finally that if the inequalities have *no* solution, the agenda has *no* consistent quota rule. This is often so for agendas in classical logic, since here the judgments on atomic propositions fully settle the judgments on compound propositions.

While theoretically elegant, Theorem 3's system of inequalities is abstract. Checking whether it holds requires to know which sets are irreducible. The latter question can even be *non-decidable* in the technical sense: in some logics (such as standard predicate logic), it is non-decidable whether a set of propositions is inconsistent; so derived notions like irreducibility or minimal inconsistency may also be non-decidable.

In view of applications, two corollaries are useful. Call an inconsistent set *trivial* if it contains a pair $p, \neg p$ or contains a contradiction p (like $a \wedge \neg a$). Any trivial $Y \in \mathcal{IR}_<$ has by minimal inconsistency the form $Y = \{p, \neg p\}$ or $Y = \{p\}$. So for trivial $Y \in \mathcal{IR}_<$ the inequality $\sum_{p \in Y}(n - m_p) < n$ holds automatically, whatever the thresholds $(m_p)_{p \in X^+} \in \{1, ..., n\}^{X^+}$. Removing these redundant inequalities, we obtain a slightly slimmer characterisation:

Corollary 4 *Theorem* 3 *still holds if* $\mathcal{IR}_<$ *is replaced by* $\mathcal{IR}^*_< := \{Y \in \mathcal{IR}_< : Y$ *is non-trivial*$\}$.

As an illustration, consider a simple implication agenda X. By Theorem 1, $F_{(m_p)_{p \in X^+}}$ is consistent if and only if

$$m_b \leq m_a + m_{a \to b} - n \text{ for all } a \to b \in X. \tag{19}$$

This characterisation is equivalent to that of Corollary 4 if $<$ is defined by $Z < Y :\Leftrightarrow |Z| < |Y|$: indeed, $\mathcal{IR}^*_< = \{\{a, a \to b, \neg b\} : a \to b \in X\}$ by Example 2 above, so that

[27] If we endow each $p \in X^+$ not with a threshold $m_p \in \{1, ..., n\}$ but, more generally, with a set $\mathcal{C}_p$ of ("winning") coalitions $C \subseteq N$ such that $\emptyset \notin \mathcal{C}_p$, $N \in \mathcal{C}_p$, and $[C \in \mathcal{C}_p \& C \subseteq C^* \subseteq N] \Rightarrow C^* \in \mathcal{C}_p$, we can define an aggregation rule $F_{(\mathcal{C}_p)_{p \in X^+}}$ with universal domain by $F_{(\mathcal{C}_p)_{p \in X^+}}(A_1, ..., A_n) = \{p \in X : \{i \in N : p \in A_i\} \in \mathcal{C}_p\}$ (where $\mathcal{C}_{\neg p} := \{C \subseteq N : N \backslash C \notin \mathcal{C}_p\}$ for all $p \in X^+$). Such a rule $F_{(\mathcal{C}_p)_{p \in X^+}}$ is called a *committee rule*. The quota rules $F_{(m_p)_{p \in X^+}}$ are precisely the anonymous committee rules (where each $p \in X^+$ has set of winning coalitions $\mathcal{C}_p = \{C \subseteq N : |C| \geq m_p\}$). The analogue of Theorem 3 is: for any simplicty relation $<$, a committee rule $F_{(\mathcal{C}_p)_{p \in X^+}}$ is consistent if and only if $\cap_{p \in Y} C_p \neq \emptyset$ for all $Y \in \mathcal{IR}_<$ and all winning coalitions $C_p \in \mathcal{C}_p$, $p \in Y$. This becomes the non-anonymous intersection property result if $\mathcal{IR}_< = \mathcal{MI}$, i.e. if we choose $< := \subsetneq_{\text{finite}}$.

[28] More precisely, it does not require for propositions $p \in X$ that $m_{\neg p} = n - m_p + 1$ (or, in the non-anonymous case discussed in footnote 27, that the coalitions winning for $\neg p$ be the coalitions whose complements are not winning for p).

the inequalities $\sum_{p \in Y}(n - m_p) < n$, $Y \in \mathcal{IR}^*_<$, are equivalent to the inequalities (19). If $<$ is alternatively defined as $\subsetneq_{\text{finite}}$, then by Example 1 above $\mathcal{IR}^*_<$ consists of all sets of type (17); thus $\mathcal{IR}^*_<$ is now much larger, and the resulting characterisation of Corollary 4 contains redundant inequalities.

Determining the set $\mathcal{IR}_<$ (or $\mathcal{IR}^*_<$) is often hard, e.g. for general implication agendas. Determining a *superset* of it can be simpler – and it suffices by the next corollary, obtained by combining Corollary 4 with Theorem 3, the latter applied with $< = \subsetneq_{\text{finite}}$, i.e. with $\mathcal{IR}_< = \mathcal{MI}$.

Corollary 5 *Theorem* 3 *still holds if* $\mathcal{IR}_<$ *is replaced by any* $\mathcal{Y}$ *with* $\mathcal{IR}^*_< \subseteq \mathcal{Y} \subseteq \mathcal{MI}$.

So, to find out for a concrete agenda which quota rules are consistent, it suffices to define a suitable simplicity relation $<$ and determine *some* set $\mathcal{Y}$ with $\mathcal{IR}^*_< \subseteq \mathcal{Y} \subseteq \mathcal{MI}$. Precisely this is done for implication agendas in the appendix.

8 Conclusion

Connection rules, of the uni-directional kind $p \to q$ or bi-directional kind $p \leftrightarrow q$, are at the heart of judgment aggregation. They express links that may be accepted or rejected, for instance causal links between facts or justificational links between claims. Once we interpret these (bi-)implications subjunctively, we can generate consistent and complete collective judgment sets by taking independent and anonymous votes on the propositions, provided that we use appropriate acceptance thresholds (see Theorems 1 and 2 and Table 2). This possibility result holds for judgment aggregation problems on so-called *implication* agendas.

The results on implication agendas are applications of an abstract result, Theorem 3, which applies to arbitrary agendas in a general logic: it characterises consistent aggregation in terms of so-called *irreducible* sets (which generalise minimal inconsistent sets[29]). It would be interesting to apply this result to classes of agendas other than implication agendas, in order to gain new insights on (im)possibilities of propositionwise voting. However, at least as important as this would be to develop a systematic understanding of *non-propositionwise* judgment aggregation rules. Though often mentioned, this route is largely unexplored.

9 References

[1] B. Chapman, Rational aggregation, Polit. Philos. Econ. 1(3) (2002), 337-354.

[2] F. Dietrich, Judgment aggregation: (im)possibility theorems, J. Econ. Theory 126(1) (2006), 286-298.

[3] F. Dietrich, A generalised model of judgment aggregation, Soc. Choice Welfare 28 (2007), 529-65.

[4] F. Dietrich, Aggregation theory and the relevance of some issues to others, Meteor Research Memorandum 07/002 (2006).

[29] Because minimal inconsistent sets are defined relative to set-inclusion whereas irreducible sets are defined relative to a general *simplicity* relation that need not be set-inclusion

[5] F. Dietrich, C. List, Strategy-proof judgment aggregation, Econ. Philos. 23 (2007), 269-300.
[6] F. Dietrich, C. List, Judgment aggregation by quota rules, J. Theoretical Politics 19(4) (2007), 391-424.
[7] F. Dietrich, C. List, Arrow's theorem in judgment aggregation, Soc. Choice Welfare 29 (2007), 19-33.
[8] F. Dietrich, C. List, Judgment aggregation on restricted domains, Meteor Research Memorandum 06/033 (2006).
[9] E. Dokow, R. Holzman, Aggregation of binary evaluations, working paper (2005), Technion Israel Institute of Technology.
[10] P. Gärdenfors, An Arrow-like theorem for voting with logical consequences, Econ. Philos. 22(2) (2006), 181-190.
[11] S. Konieczny, R. Pino-Perez, Merging information under constraints: a logical framework, Journal of Logic and Computation 12(5) (2002), 773-808.
[12] L.A. Kornhauser, L.G. Sager, Unpacking the Court. Yale Law Journal 96(1) (1986), 82-117.
[13] D. Lewis, Counterfactuals, Oxford: Basil Blackwell, 1973.
[14] C. List, A model of path dependence in decisions over multiple propositions, Amer. Polit. Sci. Rev. 98(3) (2004), 495-513.
[15] C. List, P. Pettit, Aggregating sets of judgments: an impossibility result, Econ. Philos. 18 (2002), 89-110.
[16] C. List, P. Pettit, Aggregating sets of judgments: two impossibility results compared, Synthese 140(1-2) (2004), 207-235.
[17] P. Mongin, Factoring out the impossibility of logical aggregation, J. Econ. Theory, forthcoming.
[18] K. Nehring, C. Puppe, Strategyproof social choice on single-peaked domains: possibility, impossibility and the space between, working paper (2002), Karlsruhe University.
[19] K. Nehring, C. Puppe, Consistent judgment aggregation: the truth-functional case, Social Choice and Welfare, forthcoming.
[20] K. Nehring, C. Puppe, The structure of strategy-proof social choice, part II: non-dictatorship, anonymity and neutrality, working paper (2005), Karlsruhe University.
[21] M. Pauly, M. van Hees, Logical constraints on judgment aggregation, J Philos Logic 35 (2006), 569-585.
[22] P. Pettit, Deliberative democracy and the discursive dilemma, Philosophical Issues 11 (2001), 268-299.
[23] G. Pigozzi, Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation, Synthese 152(2) (2006), 285-298.
[24] G. Priest, An introduction to non-classical logic, Cambridge University Press, 2001
[25] R. Stalnaker, A theory of conditionals, in N. Rescher (Ed.) Studies in Logical Theory, Blackwell: Oxford, 1986.
[26] M. van Hees, The limits of epistemic democracy, Soc. Choice Welfare 28 (2007), 649-66.

A Proof of Theorem 2 from Theorem 3

We consider an arbitrary implication agenda X ($\subseteq \mathbf{L}$). The language $\mathbf{L}$ (defined in Section 2) is endowed with the non-classical notions of entailment and (in)consistency defined in Section 3 (using C^+-interpretations). Recall that $\mathcal{A}$ ($\subseteq \mathbf{L}$) is the set of atomic propositions. Denote the set of all connection rules by $\mathcal{R}$ ($= \{a \to b, (a \wedge b) \leftrightarrow c, ...\}$). For all $S \subseteq \mathbf{L}$ let $S^{\neg} := \{\neg p : p \in S\}$ and $\bar{S} := S \cup S^{\neg}$. We wish to apply Theorem 3 to X – but with which simplicity relation $<$? Defining $<$ as $\subsetneq_{\text{finite}}$ gives a very complicated set $\mathcal{IR}_< = \mathcal{MI}$ (containing for instance sets like $Y = \{a, a \to b, a', a' \to b', (b \wedge b') \to (a \wedge c), \neg c\}$). Even the finer simplicity relation given by $Z < Y :\Leftrightarrow |Z| < |Y|$, while suitable for *simple* implication agendas (see the end of Section 7), is inappropriate in general since, as indicated in Section 7, we would like to simplify sets like $\{a, a \to (b_1 \wedge b_2 \wedge b_3), (b_1 \wedge b_2 \wedge b_3) \to c, \neg c\}$ into $\{b_1, b_2, b_3, (b_1 \wedge b_2 \wedge b_3) \to c, \neg c\}$ on the grounds that the latter set contains fewer connection rules despite of having more elements overall. The following lexicographic simplicity notion allows us to perform such simplifications and to get a grip on $\mathcal{IR}_<$. For all inconsistent sets $Z, Y \subseteq X$,

$$Z < Y :\Leftrightarrow (|Z \cap \bar{\mathcal{R}}|, |Z \cap \bar{\mathcal{A}}|) \text{ is lexicographically smaller than } (|Y \cap \bar{\mathcal{R}}|, |Y \cap \bar{\mathcal{A}}|),$$

i.e. $|Z \cap \bar{\mathcal{R}}| < |Y \cap \bar{\mathcal{R}}|$ or $|Z \cap \bar{\mathcal{R}}| = |Y \cap \bar{\mathcal{R}}| \& |Z \cap \bar{\mathcal{A}}| < |Z \cap \bar{\mathcal{A}}|$. For instance, $\{a, \neg b\} < \{a \to b\}$ as $(0, 2)$ is lexicographically smaller than $(1, 0)$.

The following is easily shown (using that the "lexicographically smaller than" relation is well-founded).

Lemma 6 *The above relation $<$ is a simplicity relation.*[30]

To identify the $<$-irreducible sets, we first need to understand better which entailments and inconsistencies hold within implication agendas; hence the next two technical lemmas. Generalising Section 6's notation "X_p", I put, for all $p \in \mathbf{L}$ and all $R \subseteq \mathbf{L}$,

$$R_p := \{s \in \mathbf{L} : p \to s \in R \text{ or } p \leftrightarrow s \in R \text{ or } s \leftrightarrow p \in R\},$$

the set of propositions "reachable" from p via (bi-)implications in R. I first establish a plausible fact about entailments between connection rules: namely, for instance, that $R = \{p \to b, p \to (c \wedge d)\} \vDash p \to (b \wedge c)$ because each conjunct of $b \wedge c$ (i.e. b and c) is a conjunct of some $s \in R_p = \{b, c \wedge d\}$.

Lemma 7 *For all $R \subseteq \mathcal{R}$ and $p \to q \in \mathcal{R}$,*

$$R \vDash p \to q \Leftrightarrow \mathrm{C}(q) \backslash \mathrm{C}(p) \subseteq \bigcup_{s \in R_p} \mathrm{C}(s).$$

Note that this characterisation of $R \vDash p \to q$ implies one of $R \vDash p \leftrightarrow q$ (for $R \subseteq \mathcal{R}$ and $p \leftrightarrow q \in \mathcal{R}$), since $R \vDash p \leftrightarrow q$ if and only if $R \vDash p \to q$ and $R \vDash q \to p$.

Proof. Let $R \subseteq \mathcal{R}$ and $p \to q \in \mathcal{R}$.

[30] More generally, for any partition of X into sets $X_1, ..., X_k$, a simplicity relation $<$ is defined by $Z < Y :\Leftrightarrow [(|Z \cap X_1|, ..., |Z \cap X_n|)$ is lexicographically smaller than $(|Y \cap X_1|, ..., |Y \cap X_n|)]$.

1. First let $\mathrm{C}(q)\backslash\mathrm{C}(p) \subseteq \bigcup_{s\in R_p} \mathrm{C}(s)$. Suppose all $r \in R$ hold in world w of interpretation $(W,(f_r),(v_w))$. We have to show that $p \to q$ holds in w, i.e. that all $a \in \mathrm{C}(q)$ hold in all $w^* \in f_p(w)$. Let $a \in \mathrm{C}(q)$ and $w^* \in f_p(w)$. By assumption, $a \in \mathrm{C}(p)$ or $a \in \mathrm{C}(s)$ for some $s \in R_p$. In the first case, a holds in w^* as p does (by $w^* \in f_p(w)$). In the second case, a holds in w^* as s does (by $v_w(p \to s) = T$ and $w^* \in f_p(w)$).

2. Conversely, suppose that $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$ but $a \notin \bigcup_{s\in R_p} \mathrm{C}(s)$. To show $R \nvDash p \to q$, consider an interpretation $(W,(f_p),(v_w))$ such that: (i) W contains at least two distinct worlds w, w^*, (ii) all atomic propositions hold in w, (iii) all atomic propositions except a hold in w^*, (iv) $f_p(w) = \{w, w^*\}$ (which is allowed as p holds in w and w^*), and (v) for all $t \in \mathbf{L}\backslash\{p\}$ $f_t(w) \subseteq \{w\}$. To complete the proof, I show that all $r \in R$ hold in w but $p \to q$ doesn't. First, $v_w(p \to q) = F$ by (iv) and as $v_{w^*}(q) = F$ by (iii). To show the truth in w of all $r \in R$, I show that of every implication $t \to s$ with $t \to s \in R$ or $t \leftrightarrow s \in R$ or $s \leftrightarrow t \in R$. For such $t \to s$, if $t \neq p$ then $v_w(t \to s) = T$ by (v) and (ii); and if $t = p$ then $v_w(t \to s) = T$ by (iv) and (ii)-(iii) and using that $a \notin \mathrm{C}(s)$. ■

The next technical lemma shows that there are broadly two ways in which a subset A of the implication agenda X can be inconsistent (the second way, (20), holds for instance if $\neg(a \to (b \wedge c)), a \to b, a \to c \in A$.).

Lemma 8 *If $A \subseteq \bar{\mathcal{A}} \cup \bar{\mathcal{R}}$ is inconsistent, then either already $A\backslash\mathcal{R}^{\neg}$ is inconsistent or*

$$A \text{ contains some } \neg r \in \mathcal{R}^{\neg} \text{ such that } A \cap \mathcal{R} \vDash r. \tag{20}$$

Proof. Suppose $A \subseteq \bar{\mathcal{A}} \cup \bar{\mathcal{R}}$. Assume $A_* := A\backslash\mathcal{R}^{\neg}$ is consistent and (20) does not hold. I show that A is consistent. For all $\neg(p \to q) \in A$,

(α) there is $a_{p\to q} \in \mathrm{C}(q)\backslash\mathrm{C}(p)$ with $a_{p\to q} \notin \mathrm{C}(q')$ for all $q' \in A_p$,

as otherwise $\mathrm{C}(q)\backslash\mathrm{C}(p) \subseteq \cup_{q'\in A_p}\mathrm{C}(q')$, whence by Lemma 7 $A \cap \mathcal{R} \vDash p \to q$ (take $R := A \cap \mathcal{R}$ and note that $R_p = A_p$), implying (20). Further, for all $\neg(p \leftrightarrow q) \in A$, *either*

($\beta1$) there is $a^1_{p\leftrightarrow q} \in \mathrm{C}(q)\backslash\mathrm{C}(p)$ with $a^1_{p\leftrightarrow q} \notin \mathrm{C}(q')$ for all $q' \in A_p$

or

($\beta2$) there is $a^2_{p\leftrightarrow q} \in \mathrm{C}(p)\backslash\mathrm{C}(q)$ with $a^2_{p\leftrightarrow q} \notin \mathrm{C}(p')$ for all $p' \in A_q$,

as otherwise $\mathrm{C}(q)\backslash\mathrm{C}(p) \subseteq \cup_{q'\in A_p}\mathrm{C}(q')$ and $\mathrm{C}(p)|\mathrm{C}(q) \subseteq \cup_{p'\in A_q}$, whence again by Lemma 7 $A \cap \mathcal{R} \vDash p \to q$ and $A \cap \mathcal{R} \vDash p \to r$, i.e. $A \cap \mathcal{R} \vDash p \leftrightarrow q$, implying (20).

To prove A's consistency, I construct an interpretation and show that in a world all $r \in A$ hold. Notationally, for any $r \in \mathcal{R}$ let r^{mat} be r's material counterpart: $(p \to q)^{\mathrm{mat}}$ is $\neg p \vee q$, and $(p \leftrightarrow q)^{\mathrm{mat}}$ is $(p \to q)^{\mathrm{mat}} \wedge (q \to p)^{\mathrm{mat}}$. Let A_*^{mat} be the set arising from A_* by replacing all $r \in A_* \cap \mathcal{R}$ by r^{mat}. Since A_* is consistent and $r \models r^{\mathrm{mat}}$ for all $r \in \mathcal{R}$, A_*^{mat} is also consistent. So there exists an interpretation $(W,(f_p),(v_w))$ and a world w such that

(w1) all members of A_*^{mat} are true in w.

As the propositions in A_*^{mat} contain no subjunctive (bi-)implications, their truth values in w depend neither on other worlds nor on the functions $f_p, p \in \mathbf{L}$. So we may assume the following w.l.o.g.

(w2) For all $\neg(p \to q) \in A$, there is a world $w_{p\to q} \in W\backslash\{w\}$ in which all atomic proposition except $a_{p\to q}$ hold; and $w_{p\to q} \in f_p(w)$ but $w_{p\leftrightarrow q} \notin f_s(w)$ $\forall s \in \mathbf{L}\backslash\{p\}$.

(w3) For all $\neg(p \leftrightarrow q) \in A$ with (β1), there is a world $w^1_{p\leftrightarrow q} \in W\backslash\{w\}$ in which all atomic propositions except $a^1_{p\leftrightarrow q}$ hold; and $w^1_{p\leftrightarrow q} \in f_p(w)$ but $w^1_{p\leftrightarrow q} \notin f_s(w)$ $\forall s \in \mathbf{L}\backslash\{p\}$.

(w4) For all $\neg(p \leftrightarrow q) \in A$ with (β2), there is a world $w^2_{p\leftrightarrow q} \in W\backslash\{w\}$ in which all atomic propositions except $a^2_{p\leftrightarrow q}$ hold; and $w^2_{p\leftrightarrow q} \in f_q(w)$ but $w^2_{p\leftrightarrow q} \notin f_s(w)$ $\forall s \in \mathbf{L}\backslash\{q\}$.

(w5) Worlds $w' \in W$ other than those defined in (w1)-(w4) are not reachable from w: $w' \notin f_r(w)$ $\forall r \in \mathbf{L}$.

To complete the proof, I consider any $r \in A$ and show that r holds in w.

Case 1: r is atomic or negated atomic. Then $r \in A^{\text{mat}}_*$. So r holds in w by (w1).

Case 2: r is an implication $s \to t$. Let $w' \in f_s(w)$. I have to show that t holds in w'. If $w' = w$, s holds in w by $w \in f_s(w)$; so, as $(s \to t)^{\text{mat}} = \neg s \vee t$ holds in w by (w1), t holds in w. Now let $w' \neq w$. Then by (w5), w' is one of the worlds defined in (w2)-(w4). Assume $w' = w_{p\to q}$, a world defined in (w2) (proofs for (w3) and (w4) are similar). By $w_{p\to q} \in f_s(w)$ and (w2), $p = s$. By (w2), all atomic propositions except $a_{p\to q}$ hold in $w_{p\to q}$, where $a_{p\to q}$ isn't a conjunct of t by (α). So t holds in $w_{p\to q} = w'$.

Case 3: r is a bi-implication $s \leftrightarrow t$. $s \leftrightarrow t$ holds in w if $s \to t$ and $t \to s$ are true in w. The latter can be shown by a procedure analogous to that in case 2.

Case 4: r is a negated implication $\neg(p \to q)$. To show that r holds in w, I show that $p \to q$ fails in w. This is so because, by (w2), $w_{p\to q} \in f_p(w)$ where q fails in $w_{p\to q}$ as its conjunct $a_{p\to q}$ fails.

Case 5: r is a negated bi-implication $\neg(p \leftrightarrow q)$. To show that r holds in w, I show that $p \leftrightarrow q$ is false in w, i.e. that $p \to q$ or $q \to p$ is false in w. Under (β1) $p \to q$ is false in w (consider the world $w^1_{p\leftrightarrow q}$ and use (w3)), and under (β2) $q \to p$ is false in w (consider the world $w^2_{p\leftrightarrow q}$ and use (w4)). ■

To allow us to apply Corollary 5, I now define a class $\mathcal{Y}$ of inconsistent sets $Y \subseteq X$, and I show that $\mathcal{IR}^*_< \subseteq \mathcal{Y} \subseteq \mathcal{MI}$. Let $\mathcal{Y} := \mathcal{Y}_\to \cup \mathcal{Y}_\leftrightarrow \cup \mathcal{Y}_{\neg\to} \cup \mathcal{Y}_{\neg\leftrightarrow}$, where $\mathcal{Y}_\to$, $\mathcal{Y}_\leftrightarrow$, $\mathcal{Y}_{\neg\to}$ and $\mathcal{Y}_{\neg\leftrightarrow}$ are the sets that consist, respectively, of

- all $Y \subseteq X$ of type $\{p \to q, \neg a\} \cup \mathrm{C}(p)$ where $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$;
- all $Y \subseteq X$ of type $\{p \leftrightarrow q, \neg a\} \cup \mathrm{C}(p)$ or $\{q \leftrightarrow p, \neg a\} \cup \mathrm{C}(p)$ where $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$;
- all $Y \subseteq X$ of type $\{\neg(p \to q)\} \cup \{p_s : s \in S\}$ where $S \in X_{p\to q}$ and $\forall s \in S$ $p_s \in \{p \to s, p \leftrightarrow s, s \leftrightarrow p\}$;
- all $Y \subseteq X$ of type $\{\neg(p \leftrightarrow q)\} \cup \{p_s : s \in S\} \cup \{q_s : s \in S'\}$ where $S \in X_{p\to q}$, $\forall s \in S$ $p_s \in \{p \to s, p \leftrightarrow s, s \leftrightarrow p\}$, $S' \in X_{q\to p}$, $\forall s \in S'$ $q_s \in \{q \to s, q \leftrightarrow s, s \leftrightarrow q\}$, and the sets $\{p_s : s \in S\}$, $\{q_s : s \in S'\}$ are either each disjoint with $\{p \leftrightarrow q, q \leftrightarrow p\}$ or each equal to $\{q \leftrightarrow p\}$ (the latter is only possible if $S = \{q\} \& S' = \{p\}$; the former holds automatically if $S \neq \{q\} \& S' \neq \{p\}$ as then $q \notin S \& p \notin S'$).

(The set $X_{p\to q}$ in the last two bullet points was defined in Section 6.)

Lemma 9 *For $\mathcal{Y}$ as defined above, $\mathcal{Y} \subseteq \mathcal{MI}$.*

Proof. Let $\mathcal{Y}$ be as specified. Consider any $Y \in \mathcal{Y}$. I show that $Y \in \mathcal{MI}$ by going through the four possible cases.

1. Let $Y \in \mathcal{Y}_\to$, i.e. $Y = \{p \to q, \neg a\} \cup \mathrm{C}(p)$ where $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$. Y is inconsistent because, by $\mathrm{C}(p) \vDash p$ and $\{p \to q, p\} \vDash q$, we have $\{p \to q\} \cup \mathrm{C}(p) \vDash q$.

Moreover, for any $y \in Y$, the consistency of $Y \backslash \{y\}$ can be checked by finding an interpretation with a world w in which all $z \in Y \backslash \{y\}$ hold. Specifically, $\{p \rightarrow q\} \cup \mathrm{C}(p)$ is consistent: let all atomic propositions hold in w and in all other worlds; $\{\neg a\} \cup \mathrm{C}(p)$ is consistent: let all atomic propositions except a hold in w; and, for any $y \in \mathrm{C}(p)$, $\{p \rightarrow q, \neg a\} \cup \mathrm{C}(p) \backslash \{y\}$ is consistent: let the only atomic propositions true in w be those in $\mathrm{C}(p) \backslash \{y\}$, and put $f_p(w) = \emptyset$ (which is allowed as p fails in w).

2. If $Y \in \mathcal{Y}_{\leftrightarrow}$ then $Y \in \mathcal{MI}$ by a proof similar to that under 1.

3. Now let $Y \in \mathcal{Y}_{\neg\rightarrow}$, say (in the earlier notation) $Y = \{\neg(p \rightarrow q)\} \cup \{p_s : s \in S\}$. We have $\{p_s : s \in S\}_p = S \in X_{p \rightarrow q}$, whence by Lemma 7 $\{p_s : s \in S\} \vDash p \rightarrow q$. So Y is inconsistent. To check *minimal* inconsistency, consider any $Z \subsetneq Y$. If $\neg(p \rightarrow q) \notin Z$, Z is consistent, as seen from an interpretation such that all atomic propositions hold in all worlds. If $\neg(p \rightarrow q) \in Z$, then $Z = \{\neg p \rightarrow q)\} \cup R^*$ with $R^* = \{p_s : s \in S^*\}$ and $S^* \subsetneq S$. Note that $R^*_p = S^*$. So $R^*_p \subsetneq S$. This and $S \in X_{p \rightarrow q}$ imply that $\mathrm{C}(q) \backslash \mathrm{C}(p) \nsubseteq \cup_{s \in R^*_p} \mathrm{C}(s)\}$, whence by Lemma 7 $R^* \nvDash p \rightarrow q$. So Z $(= \{\neg p \rightarrow q)\} \cup R^*)$ is consistent.

4. Finally, let $Y \in \mathcal{Y}_{\neg\leftrightarrow}$, say (in the earlier notation) $Y = \{\neg(p \leftrightarrow q)\} \cup \{p_s : s \in S\} \cup \{q_s : s \in S'\}$. It can be shown like under 3 that $\{p_s : s \in S\} \vDash p \rightarrow q$ and $\{q_s : s \in S'\} \vDash q \rightarrow p$. So $\{p_s : s \in S\} \cup \{q_s : s \in S'\} \vDash p \leftrightarrow q$. Hence Y is inconsistent. Now consider any $Z \subsetneq Y$. If $\neg(p \leftrightarrow q) \notin Z$, Y is consistent by an argument like in case 3. If $\neg(p \leftrightarrow q) \in Z$, then $Z = \{\neg(p \leftrightarrow q)\} \cup R^*$ with $R^* = \{p_s : s \in S^*\} \cup \{q_s : s \in S'^*\}$ and $S^* \subseteq S$, $S'^* \subseteq S'$, where $S^* \subsetneq S$ or $S'^* \subsetneq S'$. Note that $R^*_p = S^*$ and $R^*_q = S'^*$. So $R^*_p \subsetneq S$ or $R^*_q \subsetneq S'$. Hence $R^* \nvDash p \rightarrow q$ or $R^* \nvDash q \rightarrow p$, by an argument like that under 3. So $R^* \nvDash p \leftrightarrow q$. Hence Z $(= \{\neg(p \leftrightarrow q)\} \cup R^*)$ is consistent. ■

Lemma 10 *For $<$ and $\mathcal{Y}$ as defined above, $\mathcal{IR}^*_< \subseteq \mathcal{Y}$.*

Proof. Let $<$ and $\mathcal{Y}$ $(= \mathcal{Y}_{\rightarrow} \cup \mathcal{Y}_{\leftrightarrow} \cup \mathcal{Y}_{\neg\rightarrow} \cup \mathcal{Y}_{\neg\leftrightarrow})$ be as specified. Consider a $Y \in \mathcal{IR}^*_<$. I show that $Y \in \mathcal{Y}$. I will use that $Y \in \mathcal{MI}$ by Lemma 4, and that (*) Y contains no pair $t, \neg t$ by non-triviality.

Case 1: $Y \cap \mathcal{R}^\neg = \emptyset$. Then (i) Y has a subset of type $\{p \rightarrow q\} \cup \mathrm{C}(p)$, *or* (ii) Y has a subset of type $\{p \leftrightarrow q\} \cup \mathrm{C}(p)$ or $\{q \leftrightarrow p\} \cup \mathrm{C}(p)$. Otherwise Y would be consistent, as seen from an interpretation with a world w in which the only true atomic propositions are those in Y and such that $f_t(w) = \emptyset$ if $t \in \mathbf{L}$ is false in w: in w, all $y \in Y \cap \bar{\mathcal{A}}$ hold by construction (and by (*)), all $p \rightarrow q \in Y$ hold by $f_p(w) = \emptyset$ (as p is false by not-(i)), all $p \leftrightarrow q \in Y$ hold by $f_p(w) = f_q(w) = \emptyset$ (as p and q are false by not-(ii)), and there are no $y \in Y \cap \mathcal{R}^\neg$.

Subcase 1a: (i) holds, say $\{p \rightarrow q\} \cup \mathrm{C}(p) \subseteq Y$. I show that $Y \in \mathcal{Y}_{\rightarrow}$. If there is an $a \in \mathrm{C}(q) \backslash \mathrm{C}(p)$ with $\neg a \in Y$, then $\{p \rightarrow q, \neg a\} \cup \mathrm{C}(p) \subseteq Y$, hence $\{p \rightarrow q, \neg a\} \cup \mathrm{C}(p) = Y$ (as $Y \in \mathcal{MI}$), and so $Y \in \mathcal{Y}_{\rightarrow}$. Hence it suffices to prove that such an a exists. For a contradiction, suppose (**) $\neg a \notin Y$ for all $a \in \mathrm{C}(q) \backslash \mathrm{C}(p)$. I show that Y is reducible to $Z := Y \cup \mathrm{C}(q) \backslash \{p \rightarrow q\}$, a contradiction. First, Z is indeed inconsistent: otherwise there would exist an interpretation with a world w in which all $z \in Z$ hold, where by $Z \cap \mathcal{R}^\neg = \emptyset$ we may assume w.l.o.g. that $f_p(w)$ contains no world other than w; thus $p \rightarrow q$ also holds in w, so that $Z \cup \{p \rightarrow q\} = Y \cup \mathrm{C}(q)$ is consistent, a contradiction. Second, we have $Z < Y$ by $|Z \cap \bar{\mathcal{R}}| = |Y \cap \bar{\mathcal{R}}| - 1$ (and by our lexicographic definition of $<$). Finally, any $y \in Z \backslash Y$ belongs to $\mathrm{C}(q)$, hence is entailed by $Z := \mathrm{C}(p) \cup \{p \rightarrow q\}$ $(\subseteq Y)$; it remains to show $Z \cup \{\neg y\} < Y$, which I do by proving that $|(Z \cup \{\neg y\}| \cap \bar{\mathcal{R}}| < |Y \cap \bar{\mathcal{R}}|$, i.e. that $|Y \cap \bar{\mathcal{R}}| > 1$. Suppose the contrary. Then $Y = \{p \to q\} \cup \mathrm{C}(p) \cup Y'$ for some $Y' \subseteq \bar{\mathcal{A}}$. By $Y \in \mathcal{MI}$, $\mathrm{C}(p) \cup Y'$ is consistent. So there is an interpretation with a world w in which all $a \in \mathrm{C}(p) \cup Y'$ hold, where by $\mathrm{C}(p) \cup Y' \subseteq \bar{\mathcal{A}}$ we may assume w.l.o.g. that $f_p(p)$ contains no world other than w, and that all $a \in \mathcal{A}$ with $\neg a \notin Y$ hold in w. All $a \in \mathrm{C}(q)$ satisfy $\neg a \notin Y$: if $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$ by (**), and if $a \in \mathrm{C}(q) \cap \mathrm{C}(p)$ by (*). So, in w, all $a \in \mathrm{C}(q)$ and hence q hold; so $p \to q$ holds. But then all $y \in Y$ hold in w, contradicting Y's inconsistency.

Subcase 1b: (ii) holds, say $\{p \leftrightarrow q\} \cup \mathrm{C}(p) \subseteq Y$ (the proof is analogous if $p \leftrightarrow q$ is replaced by $q \leftrightarrow p$). To show that $Y \in \mathcal{Y}_{\leftrightarrow}$, it suffices to slightly adapt the proof in Subcase 1a: replace "$\to$" by "$\leftrightarrow$", and in both interpretations assume w.l.o.g. that $f_q(w)$ (in addition to $f_p(w)$) contains no world other than w.

Case 2: $Y \cap \mathcal{R}^{\neg} \neq \emptyset$. Then $Y\backslash\mathcal{R}^{\neg} \subsetneq Y$, whence $Y\backslash\mathcal{R}^{\neg}$ is consistent by $Y \in \mathcal{MI}$. So by Lemma 8 Y contains a $\neg r \in \mathcal{R}^{\neg}$ such that $Y \cap \mathcal{R} \vDash r$. Let $R := Y \cap \mathcal{R}$. As Y is *minimal* inconsistent, $Y = \{\neg r\} \cup R$. I consider two subcases.

Subcase 2a: r is an implication $p \to q$. I show that $Y \in \mathcal{Y}_{\neg\to}$. As $Y = \{\neg(p \to q)\} \cup R \in \mathcal{MI}$, R is minimal subject to entailing $p \to q$. So, by Lemma 7, R is minimal subject to $\mathrm{C}(q)\backslash\mathrm{C}(p) \subseteq \bigcup_{s\in R_p} \mathrm{C}(s)$. This implies that $R_p \in X_{p\to q}$ and that $R = \{p_s : s \in R_p\}$ for some $p_s \in \{p \to s, p \leftrightarrow s, s \leftrightarrow p\}$, $s \in R_p$. So Y $(= \{\neg(p \to q)\} \cup R)$ is in $\mathcal{Y}_{\neg\to}$.

Subcase 2b: r is an bi-implication $p \leftrightarrow q$. I show $Y \in \mathcal{Y}_{\neg\leftrightarrow}$. Write $R = R^1 \cup R^2 \cup T$ with $R^1 := R \cap \{p \to s, p \leftrightarrow s, s \leftrightarrow p : s \in \mathbf{L}\}$, $R^2 := R \cap \{q \to s, q \leftrightarrow s, s \leftrightarrow q : s \in \mathbf{L}\}$ and $T := R\backslash(X_p \cup X_q)$. As $Y = \{\neg(p \leftrightarrow q)\} \cup R$ is minimal inconsistent, R is minimal subject to entailing $p \leftrightarrow q$, i.e. minimal subject to entailing each of $p \to q$ and $q \to p$. So, by Lemma 7 and using that $R_p = R^1_p$ and $R_q = R^2_q$, the set R is minimal subject to satisfying both (a) $\mathrm{C}(q)\backslash\mathrm{C}(p) \subseteq \bigcup_{s\in R^1_p} \mathrm{C}(s)$ and (b) $\mathrm{C}(p)\backslash\mathrm{C}(q) \subseteq \bigcup_{s\in R^2_q} \mathrm{C}(s)$. It follows that $R = R^1 \cup R^2$ (i.e. $T = \emptyset$).

First suppose $q \leftrightarrow p \in R^1$ or $q \leftrightarrow p \in R^2$. Then $Y = \{\neg(p \leftrightarrow q)\} \cup R \supseteq \{\neg(p \leftrightarrow q), q \leftrightarrow p\}$, hence by minimal inconsistency $Y = \{\neg(p \leftrightarrow q), q \leftrightarrow p\}$. So $Y \in \mathcal{Y}_{\neg\leftrightarrow}$, as desired.

Now suppose $q \leftrightarrow p \notin R^1$ and $q \leftrightarrow p \notin R^2$. As also $p \leftrightarrow q \notin R^1$ and $p \leftrightarrow q \notin R^2$ by (*), we have $R^1 \cap R^2 = \emptyset$. This and the fact that the set $Y = R^1 \cup R^2$ is minimal subject to (a)&(b) imply that R^1 is minimal subject to (a) and that R^2 is minimal subject to (b). So (like in Subcase 2a) $R^1_p \in X_{p\to q}$ with $R^1 = \{p_s : s \in R_p\}$ for some $p_s \in \{p \to s, p \leftrightarrow s, s \leftrightarrow p\}$, $s \in R_p$, and $R^2_q \in X_{q\to p}$ with $R^2 = \{q_s : s \in R_q\}$ for some $q_s \in \{q \to s, q \leftrightarrow s, s \leftrightarrow q\}$, $s \in R_q$. So Y $(= \{\neg(p \to q)\} \cup R^1 \cup R^2)$ is in $\mathcal{Y}_{\neg\leftrightarrow}$, as desired. ■

By Lemmas 9 and 10, we can apply Corollary 5 to characterise consistent quota rules. I finally prove that this characterisation can be simplified into that in Theorem 2.

Proof of Theorem 2. Let $F_{(m_p)_{p\in X^+}}$ be a quota rule, and $\mathcal{Y}$, $\mathcal{Y}_{\to}, \mathcal{Y}_{\leftrightarrow}$, $\mathcal{Y}_{\neg\to}$, $\mathcal{Y}_{\neg\leftrightarrow}$ the sets defined above. By Corollary 5 (using Lemmas 6, 9 and 10) I have to show that (a)&(b) hold iff for all $Y \in \mathcal{Y}$ $(= \mathcal{Y}_{\to} \cup \mathcal{Y}_{\leftrightarrow} \cup \mathcal{Y}_{\neg\to} \cup \mathcal{Y}_{\neg\leftrightarrow})$

$$\sum_{y\in Y}(n - m_y) < n. \tag{21}$$

I will build up this equivalence in the following four steps.

Claim 1. The LHS inequalities in (a) hold iff (21) holds for all $Y \in \mathcal{Y}_{\rightarrow}$.

Claim 2. Given (b), the RHS inequalities in (a) hold iff (21) holds for all $Y \in \mathcal{Y}_{\neg\rightarrow}$.

Claim 3. (b) holds iff (21) holds for all $Y \in \mathcal{Y}_{\leftrightarrow}$.

By Claims 1-3, (a)&(b) hold iff (21) holds for all $Y \in \mathcal{Y}_{\rightarrow} \cup \mathcal{Y}_{\leftrightarrow} \cup \mathcal{Y}_{\neg\rightarrow}$; which is the case iff (21) holds for all $Y \in \mathcal{Y}_{\rightarrow} \cup \mathcal{Y}_{\leftrightarrow} \cup \mathcal{Y}_{\neg\rightarrow} \cup \mathcal{Y}_{\neg\leftrightarrow}$, because of our last claim which completes the proof.

Claim 4. If (21) holds for all $Y \in \mathcal{Y}_{\rightarrow} \cup \mathcal{Y}_{\leftrightarrow}$ then it holds for all $Y \in \mathcal{Y}_{\neg\leftrightarrow}$ (hence the inequalities for $Y \in \mathcal{Y}_{\neg\leftrightarrow}$ are redundant in the system).

Proof of Claim 1. The inequalities (21) for all $Y \in \mathcal{Y}_{\rightarrow}$ are given by

$$(n - m_{\neg a}) + (n - m_{p\rightarrow q}) + \sum_{a' \in \mathrm{C}(p)} (n - m_{a'}) < n \ \forall p \rightarrow q \in X \ \forall a \in \mathrm{C}(q)\backslash\mathrm{C}(p).$$

Using that $n - m_{\neg a} = m_a - 1$, these inequalities can be rewritten as

$$m_a + \sum_{a' \in \mathrm{C}(p)} (n - m_{a'}) \leq m_{p\rightarrow q} \ \forall p \rightarrow q \in X \ \forall a \in \mathrm{C}(q)\backslash\mathrm{C}(p),$$

which by taking the maximum over a is equivalent to the LHS inequalities in (a).

Proof of Claim 2. Suppose (b). The inequalities (21) for all $Y \in \mathcal{Y}_{\neg\rightarrow}$ are given by

$$n - m_{\neg(p\rightarrow q)} + \sum_{s\in S}(n - m_{p_s}) < n \quad \begin{array}{l} \forall p \rightarrow q \in X \ \forall S \in X_{p\rightarrow q} \\ \forall (p_s)_{s\in S} \in (\{p \rightarrow s, p \leftrightarrow s, s \leftrightarrow p\} \cap X)^S. \end{array}$$

These inequalities can (by $n - m_{\neg(p\rightarrow q)} = m_{p\rightarrow q} - 1$) be rewritten as

$$m_{p\rightarrow q} + \sum_{s\in S}(n - m_{p_s}) \leq n \quad \begin{array}{l} \forall p \rightarrow q \in X \ \forall S \in X_{p\rightarrow q} \\ \forall (p_s)_{s\in S} \in (\{p \rightarrow s, p \leftrightarrow s, s \leftrightarrow p\} \cap X)^S, \end{array}$$

or equivalently as

$$m_{p\rightarrow q} + \max_{(p_s)_{s\in S} \in (\{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X)^S} \sum_{s\in S}(n - m_{p_s}) \leq n \ \forall p \rightarrow q \in X \ \forall S \in X_{p\rightarrow q}. \quad (22)$$

Note that

$$\max_{(p_s)_{s\in S} \in (\{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X)^S} \sum_{s\in S}(n - m_{p_s}) = \sum_{s\in S}(n - \min_{p_s \in \{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X} m_{p_s}). \quad (23)$$

For all $s \in S$ and all $p_s \in \{p \leftrightarrow q, q \leftrightarrow p\}$ we have $m_{p_s} = n$ by (b). So, for all $s \in S$, $\min_{p_s \in \{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X} m_{p_s}$ is n if $p \rightarrow s \notin X$ and $m_{p\rightarrow s}$ if $p \rightarrow s \in X$. Hence in (23) the term $(n - \min_{p_s \in \{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X} m_{p_s})$ drops out if $p \rightarrow s \notin X$ and equals $(n - m_{p\rightarrow s})$ if $p \rightarrow s \in X$. Therefore (23) implies

$$\max_{(p_s)_{s\in S} \in (\{p\rightarrow s, p\leftrightarrow s, s\leftrightarrow p\}\cap X)^S} \sum_{s\in S}(n - m_{p_s}) = \sum_{s\in S: p\rightarrow s\in X} (n - m_{p\rightarrow s}).$$

Using this, the inequalities (22) are equivalent to

$$m_{p\rightarrow q} + \sum_{s\in S: p\rightarrow s\in X} (n - m_{p\rightarrow s}) \leq n \ \forall p \rightarrow q \in X \ \forall S \in X_{p\rightarrow q},$$

and hence, as desired, to

$$m_{p\to q} + \max_{S\in X_{p\to q}} \sum_{s\in S: p\to s\in X} (n - m_{p\to s}) \leq n \;\forall p\to q \in X.$$

Proof of Claim 3. 1. First assume (21) holds for all $Y \in \mathcal{Y}_{\leftrightarrow}$, and let $p \leftrightarrow q \in X$.

1.1. Here I show $m_{p\leftrightarrow q} = n$. As $p \leftrightarrow q$ is non-degenerate, there exist $a \in \mathrm{C}(p)\backslash\mathrm{C}(q)$ and $b \in \mathrm{C}(q)\backslash\mathrm{C}(p)$. By assumption,

$$\begin{aligned}(n - m_{p\leftrightarrow q}) + (n - m_{\neg b}) + \sum_{a'\in \mathrm{C}(p)} (n - m_{a'}) &< n,\\ (n - m_{p\leftrightarrow q}) + (n - m_{\neg a}) + \sum_{b'\in \mathrm{C}(q)} (n - m_{b'}) &< n.\end{aligned}$$

Rewriting this (by using that $n - m_{\neg s} = m_s - 1$ for all $s \in X$), we obtain

$$\begin{aligned}m_{p\leftrightarrow q} - m_b + 1 &> \sum_{a'\in \mathrm{C}(p)} (n - m_{a'}) \geq n - m_a, \qquad (24)\\ m_{p\leftrightarrow q} - m_a + 1 &> \sum_{b'\in \mathrm{C}(q)} (n - m_{b'}) \geq n - m_b.\end{aligned}$$

So

$$m_{p\leftrightarrow q} \geq n - m_a + m_b \text{ and } m_{p\leftrightarrow q} \geq n - m_b + m_a. \qquad (25)$$

Adding both inequalities, we get $2m_{p\leftrightarrow q} \geq 2n$, whence $m_{p\leftrightarrow q} = n$.

1.2. Next I show that all $a \in \mathrm{C}(p)\Delta\mathrm{C}(q)$ have the same threshold. As $\mathrm{C}(p)\Delta\mathrm{C}(q)$ is the union of the non-empty sets $\mathrm{C}(p)\backslash\mathrm{C}(q)$ and $\mathrm{C}(q)\backslash\, \mathrm{C}(p)$, it is sufficient to show that $m_a = m_b$ for all $a \in \mathrm{C}(p)\backslash\mathrm{C}(q)$ and $b \in \mathrm{C}(q)\backslash\mathrm{C}(p)$. Consider such a, b. The argument in 1.1 yields (25), which by $m_{p\leftrightarrow q} = n$ implies $m_a \geq m_b$ and $m_b \geq m_a$, whence $m_a = m_b$.

1.3. Let m be the common threshold of all $a \in \mathrm{C}(p)\Delta\mathrm{C}(q)$. I suppose $m < n$ and show that $|\mathrm{C}(p)\Delta\mathrm{C}(q)| \leq 2$. The first inequality in (24) (where $b \in \mathrm{C}(q)\Delta\mathrm{C}(p)$) implies

$$m_{p\leftrightarrow q} - m_b + 1 > \sum_{a'\in \mathrm{C}(p)\backslash\mathrm{C}(q)} (n - m_{a'}),$$

which after substituting $m_{p\leftrightarrow q} = n$ and $m_b = m_{a'} = m$ gives

$$n - m \geq |\mathrm{C}(p)\backslash\mathrm{C}(q)|(n - m), \text{ i.e. } |\mathrm{C}(p)\backslash\mathrm{C}(q)| \leq 1.$$

It can be shown similarly that $|\mathrm{C}(q)\backslash\mathrm{C}(p)| \leq 1$. So $|\mathrm{C}(p)\Delta\mathrm{C}(q)| \leq 2$.

1.4. Finally, let $a'' \in \mathrm{C}(p)\cap\mathrm{C}(q)$. I show that $m_{a''} = n$. Let a, b be as in 1.1. The first inequality in (24) implies

$$m_{p\leftrightarrow q} - m_b + 1 > (n - m_{a''}) + (n - m_a),$$

which by $m_{p\leftrightarrow q} = n$ and $m_a = m_b$ implies $1 > (n - m_{a''})$, i.e. $m_{a''} = n$.

2. Conversely, assume (b). Consider any $Y \in \mathcal{Y}_{\leftrightarrow}$, say $Y = \{r, \neg a\} \cup \mathrm{C}(p)$ where $r \in \{p \leftrightarrow q, q \leftrightarrow p\}$ and $a \in \mathrm{C}(q)\backslash\mathrm{C}(p)$, and let me show (21). Using (b) and $n - m_{\neg a} = m_a - 1$,

$$\sum_{y\in Y} (n - m_y) = m - 1 + |\mathrm{C}(p)\backslash\mathrm{C}(q)|(n - m),$$

where m denotes the common threshold of all $a' \in \mathrm{C}(p)\Delta\mathrm{C}(q)$. Note that if $|\mathrm{C}(p)\backslash\mathrm{C}(q)| \geq 2$ then $|\mathrm{C}(p)\Delta\mathrm{C}(q)| \geq 3$, hence $m = n$. So, as desired,

$$\sum_{y\in Y}(n - m_y) = \begin{cases} m - 1 + n - m < n & \text{if } |\mathrm{C}(p)\backslash\mathrm{C}(q)| = 1 \\ n - 1 + 0 < n & \text{if } |\mathrm{C}(p)\backslash\mathrm{C}(q)| \geq 2. \end{cases}$$

Proof of Claim 4. Suppose (21) holds for all $Y \in \mathcal{Y}_{\rightarrow}\cup\mathcal{Y}_{\leftrightarrow}$. Consider any $Y \in \mathcal{Y}_{\neg\leftrightarrow}$, say (in the earlier notation) $Y = \{\neg(p \leftrightarrow q)\} \cup \{p_s : s \in S\} \cup \{q_s : s \in S'\}$. To prove the corresponding inequality,

$$(n - m_{\neg(p\leftrightarrow q)}) + \sum_{s\in S}(n - m_{p_s}) + \sum_{s\in S'}(n - m_{q_s}) < n,$$

I show that $m_{p_s} = n \ \forall s \in S$ and that $m_{q_s} = n \ \forall s \in S'$; in fact, I only show the former as the latter holds analogously. Let $s \in S$. Recall that $S \in X_{p\rightarrow q}$ and $p_s \in \{p \rightarrow s, p \leftrightarrow s, s \leftrightarrow p\}$.

If $p_s \in \{p \leftrightarrow s, s \leftrightarrow p\}$ then already by Claim 3 $m_{p_s} = n$, as desired.

Now assume $p_s = p \rightarrow s$. Since $s \in S \in X_{p\rightarrow q}$, $\mathrm{C}(q)\backslash\mathrm{C}(p)$ is a subset of $\cup_{s^*\in S}\mathrm{C}(s^*)$ but not of $\cup_{s^*\in S\backslash\{s\}}\mathrm{C}(s^*)$. So there is a $b \in \mathrm{C}(s) \cap \mathrm{C}(q)\backslash\mathrm{C}(p)$. Moreover, as $p \leftrightarrow q$ is non-degenerate, there is an $a \in \mathrm{C}(p)\backslash\ \mathrm{C}(q)$. As $a, b \in \mathrm{C}(p)\Delta\mathrm{C}(q)$, we have $m_a = m_b$ by Claim 3. Using Claim 1,

$$m_{p\rightarrow q} \geq \sum_{a'\in\mathrm{C}(p)}(n - m_{a'}) + \max_{b'\in\mathrm{C}(q)\backslash\mathrm{C}(p)} m_{b'} \geq n - m_a + m_b = n,$$

whence $m_{p\rightarrow q} = n$, as desired. ■