Working paper

Aggregating Causal Judgements

Richard Bradley, Franz Dietrich and Christian List
London School of Economics

October 25, 2006

Abstract

Decision making typically requires judgements about causal relations: we need to know both the causal effects of our actions and the causal relevance of various environmental factors. Judgements about the nature and strength of causal relations often differ, even among experts. How to handle such diversity is the topic of this paper. First we consider the possibility of aggregating causal judgements via the aggregation of probabilistic ones. The broadly negative outcome of this investigation leads us to look at aggregating causal judgements independently of probabilistic ones. We do so by transcribing causal claims into the judgement aggregation framework and applying some recent results in this field. Finally we look at the implications for probability aggregation when it is constrained by prior aggregation of causal judgements.

1 Introduction

Decision making typically requires judgements about causal relations: home owners need to know whether putting locks in their doors will make their houses more secure; jurors need to know whether the accused is causally responsible for damages before they can assess whether they are legally responsible; and aid agencies need to know what the effect of spending money on different projects will be on the lives of those they are concerned about. It is also often the case that opinions about the nature and strength of causal relations differ, even among experts. How to handle such diversity of opinion is the topic of this paper. We investigate the possibility of coherently aggregating different causal judgements into a single one that may be applied to the decision problem at hand.

The basic set-up of this aggregation problem is as follows. Individuals hold diverse judgements about both the nature of the causal relations between the variables in some set $\mathbf{V} = \{V, W, ...\}$ and the probabilities of these variables taking certain values, unconditionally or conditionally on the values of other variables. The task is to construct a single aggregate judgement on the causal relations between the variables and the relevant

probabilities in a way that preserves, as much as possible, the information contained in the individuals' judgements. In doing so we assume that individuals' judgements are coherent. More generally, one might allow that individuals do not make judgements about all causal relations or all probabilities in question: their judgements could be restricted to just certain variables relevant to the decision problem at hand, or further still, to just some subset of them or just one type of judgement: causal or probabilistic.

The causal judgements of individuals could be represented in a number of different ways but here we adopt the Bayesian network framework familiar from the work of Pearl [11], Spirtes, Glymour and Scheines [12] and others in which they are represented by directed acyclic graphs (DAGs) and associated conditional probabilities. We do not intend thereby to take up a position on the nature of causal judgements, nor on the question of whether they can ultimately be analysed probabilistically.[1] Anyone who holds the view that causal judgements are just features of probability judgements – for instance that to judge that X causes Y is to hold certain conditional probabilistic judgements including that the conditional probability of X given Y exceeds its unconditional probability – is free to regard the DAG representations of an individual's causal judgement as adding no information to his or her probability judgements. We also do not think that much of what we claim depends on the DAG representation; we could just as well represent causal judgements by, for instance, counterfactual beliefs of the right kind.

A DAG represents an individual's qualitative judgement of causal relevance and irrelevance between variables. Her quantitative judgement of causal dependence is reflected in the associated conditional probabilities for the values of variables, given the values of any variables on which it is directly causally dependent. Then the individual's unconditional probabilities for the values of the variables can be computed from her unconditional probabilities for the parent variables plus these associated conditional probabilities. Consider the following example, which we will use at various points in the discussion.

Example: Predicting famine. An aid agency wishes to do some advance planning for its famine relief operations and consults a number of experts in order to determine the risk of famine in some particular region. All agree that the relevant variables are R: rainfall, Y: crop yields, P: political conflict and, of course, F: famine. But they disagree both on the causal relations between the four variables and on the probabilities of the various values the variables may take. All consider rainfall to be main determinant of crop yield but while expert 1 thinks that poor crop yield and disruptive political conflict are the main causes of famine, expert 2 thinks that the causal influence of political conflict on famine is indirect, via the effect of the disruption of agricultural production on crop yields. Expert 3 considers the relationship between political conflict and famine to be more complicated still, with political conflict both causing famine directly, by

---

[1]A probabilistic analysis may involve variables not included in the DAG we consider.

Expert 1   Expert 2   Expert 3
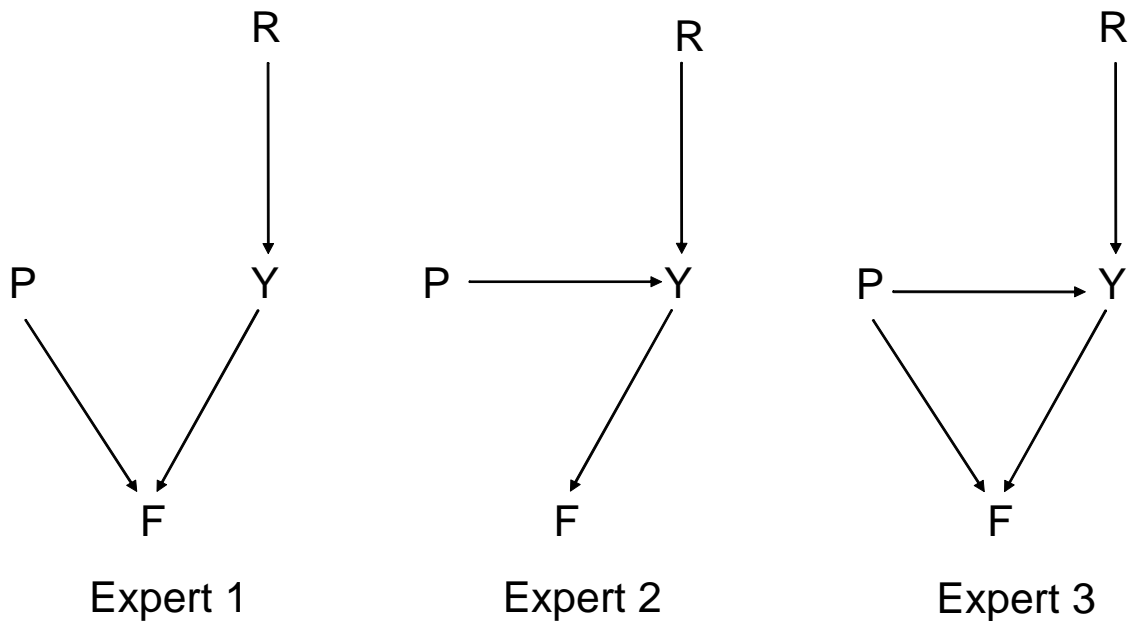
disrupting food distribution, and indirectly, through the influence on crop yields. These three opinions are represented in figure 1 by a set of DAGs.

The fact that individuals make both causal and probabilistic judgements raises the question of whether aggregation of both kinds of judgements should be conducted all at once or in two stages. In section 2, we focus on what we call *one-stage aggregation*, in which only probability judgements are aggregated. Aggregation of this sort is motivated mainly by the thought that the probability judgements of individuals reflect their causal judgements in various ways and hence that our problem may be solved by constraining probability aggregation so as to preserve the causal information contained in probability judgements. Our verdict on this possibility is broadly negative, however. In sections 3 to 5, we therefore pursue an alternative *two-stage* approach, aggregating first the causal judgements represented by the DAGs (section 3) and then the probabilistic ones (sections 4 and 5), on the assumption that a consensus about the causal relations between variables has been reached.

2   Aggregating Probability Judgements

The problem of aggregating causal judgements has not received much attention, at least in the form presented here, but there is a vast literature on aggregating expert opinion, mainly in statistics, and especially on aggregating expert probabilities. In this section, we draw on this literature to examine the possibility of reasonable one-stage aggregation

3

of individuals' judgements. One-stage aggregation may be the only method available in cases in which individuals either make no causal judgements or their causal judgements are very incomplete. It is natural moreover for those holding a probabilistic view about causation to rely only on this method. But one-stage aggregation may also be motivated by the less controversial thought that the causal judgements of individuals are reflected in (even if they are not reducible to) the relations between the individuals' unconditional and conditional probabilities for the relevant possibilities. If this is so, then even on a non-reductionist view about causal judgements one may still hope that probability aggregation could be constrained in a manner which preserved the causal judgements implicit in probabilistic ones.

Broadly, the probability aggregation literature can be divided into three approaches to the problem: linear pooling, geometric pooling and supra-Bayesian approaches. (A very useful guide to this literature can be found in the survey paper of Genest and Zidek [3]; here we draw only on salient aspects.) The last approach is directed at a slightly different problem to ours – namely that of how an individual expert should modify his judgements in the light of the expressed judgements of other experts – and so we can set it aside. The other two approaches assume that the experts' opinions have reached an equilibrium state and that no further modification of their viewpoint will take place before the relevant decision has to be made.

Consider an opinion aggregation problem of the following form. A set of *events* is given (for instance the event "high political conflict" or "low political conflict and famine"), and the task is to merge the probability judgements of individuals $1, ..., n$ (the "experts") on these events into an aggregate probability judgement on the events.[2] So, we have to merge (individual) probability functions $\mathrm{Pr}_1, ..., \mathrm{Pr}_n$ into an (aggregate) probability function Pr. Many aggregation rules are imaginable. Formally, a *(probability) aggregation rule* is a function that assigns to *any* vector $(\mathrm{Pr}_1, ..., \mathrm{Pr}_n)$ (also called a *profile*) of individual probability functions an aggregate probability function Pr.

Of the various possible aggregation rules, *linear pooling* stands out for a variety of formal and conceptual reasons (e.g., Lehrer and Wagner [5]). In particular, the following axiomatic argument can be given. Let us require the aggregation rule to satisfy two seemingly natural conditions:

IIA (*Independence of Irrelevant Alternatives*) The aggregate probability of any given event X depends only on the individuals' probabilities of X (regardless of the individuals' probabilities of other events Y).[3]

---

[2]Events can be identified with subsets of a given set of possible worlds. In many formal results, the set of events considered (i.e., the domain of the individual probability functions $\mathrm{Pr}_1, ..., \mathrm{Pr}_n$ and the aggregate probability function Pr) forms an *algebra*: the negation (complement) of any event is also an event, and the disjunction (union) of two events is an event too.

[3]Formally, $\mathrm{Pr}(X)$ is a function of $\mathrm{Pr}_1(X), ..., \mathrm{Pr}_n(X)$ (this function may be a different one for different events X).

ZP (*Zero Preservation*) The aggregate probability of any given event X is zero whenever all individuals give X zero probability.[4]

Applied to the event "famine", for instance, Zero Preservation implies that famine is assigned an aggregate probability of zero if all individuals assign a probability of zero to it. Independence of Irrelevant Alternative implies that the aggregate probability of famine depends only on the individuals' beliefs about the probability of famine, not on their beliefs about the probability of a certain level of crop yield, political conflict, etc. (This is not to deny, of course, that individuals form their beliefs regarding famine in the light of their judgements on crop yield, political conflict etc.)

Perhaps surprisingly, the *only* aggregation rules satisfying these two conditions are linear pooling functions: the aggregate probability of any event X is a (possibly weighted) arithmetic average of the individual probabilities of X, i.e.

$$\mathrm{Pr}(X) = w_1 \mathrm{Pr}_1(X) + ... + w_n \mathrm{Pr}_n(X),$$

where the weights $w_1, ..., w_n \geq 0$ add up to one and are the same for all events X.[5] Examples of linear pooling functions are equal-weight averaging ($w_1 = ... = w_n = 1/n$) and dictatorial aggregation (some individual i has weight $w_i = 1$ and all others have weight 0).

But do such linear pooling functions respect the causal knowledge of the individual experts satisfactorily? An individual's causal judgements will be reflected in certain (unconditional or conditional) independencies in her probability judgements. For instance, if individual i believes that events X and Y do not causally affect each other but have a common cause, say event Z, then he or she will take X and Y to be probabilistically independent given Z,[6] because any probabilistic correlation between X and Y is "screened off" by conditionalising on Z. A minimal condition of respecting causal judgements is that at least *unanimously* held causal judgements be reflected in the aggregate probability function Pr: i.e., Pr should display at least those (conditional) independencies required by unanimous causal judgements. For example, if *all* individuals take X and Y to be causally independent with common cause Z, then that independence judgement should be reflected in the aggregate probability function Pr. This motivates the following condition on probability aggregation:

IP (*Independence Preservation*) For any given events X, Y, Z, if all individuals i judge X and Y to be probabilistically independent given Z, then this conditional independence also holds under the aggregate probability function.[7]

---

[4]Formally, $\mathrm{Pr}(X) = 0$ if $\mathrm{Pr}_1(X) = ... = \mathrm{Pr}_n(X) = 0$.

[5]See McConway [9], drawing on previous results. The result requires that the set of events considered forms an algebra (see footnote 2) and contains at least three events apart from the contradiction (empty set of worlds) and the tautology (set of all worlds).

[6]Formally, $\mathrm{Pr}_i(XY|Z) = \mathrm{Pr}_i(X|Z)\mathrm{Pr}_i(Y|Z)$.

[7]Formally, if, for all individuals i, $\mathrm{Pr}_i(Z) > 0$ and $\mathrm{Pr}_i(XY|Z) = \mathrm{Pr}_i(X|Z)\mathrm{Pr}_i(Y|Z)$, then also $\mathrm{Pr}(Z) > 0$ and $\mathrm{Pr}(XY|Z) = \mathrm{Pr}(X|Z)\mathrm{Pr}(Y|Z)$.

Note that, by preserving all unanimous probabilistic independencies (conditional or unconditional), we may also preserve ones that are not grounded in unanimous *causal* judgements. For instance, it may be that all individuals judge X and Y to be independent given Z, but some do so on the grounds of judging that X indirectly causes Y through Z, others on the grounds of judging that Y indirectly causes X through Z, still others on the grounds of judging that X, Y and Z are entirely causally disconnected, etc. Even in this case of causal disagreement, Independence Preservation requires the preservation of the *probabilistic* (conditional) independence. The purely probabilistic informational basis of one-stage probability aggregation simply does not allow us to distinguish between different motivations (causal or other) behind probabilistic independencies. Without explicit causal information, we have to use Independence Preservation to preserve unanimous causal judgements, although this requires us to preserve even those conditional independencies that are not causally motivated.

It turns out, however, that Independence Preservation is violated by all linear pooling functions (unless some individual i gets maximal weight $w_i = 1$) and thus by all non-dictatorial probability aggregation rules satisfying the above conditions of Independence of Irrelevant Alternatives and Zero Preservation. This fact, proven in Genest and Wagner [4], can be illustrated using our earlier example. Suppose the aid agency consults a couple of experts in order to determine the risk of famine in a particular region and that both experts agree that famine is caused by a combination of drought (the event of rainfall R below some critical threshold) and political instability (the event of political conflict P above some critical threshold), which undermines local solutions to poor crop yields. Furthermore, they agree that these two factors are both causally and probabilistically independent, at least in the short term. But they disagree on the probability of drought and of political instability. Since neither speaks with greater authority than the other, the aid agency calculates its probabilities for these events by taking the average of the judgements of the two experts. Let D and I respectively denote the occurrence of drought and political instability in the region and DI their concurrence. $\Pr_1$, $\Pr_2$, and $\Pr$ are respectively the probability functions of expert 1, expert 2 and the aid agency. Since pooling happens by averaging, the aid agency will assign the following probabilities:

$$
\begin{aligned}
\Pr(D) &= \frac{\Pr_1(D) + \Pr_2(D)}{2}, \ \Pr(I) = \frac{\Pr_1(I) + \Pr_2(I)}{2}, \\
\Pr(DI) &= \frac{\Pr_1(DI) + \Pr_2(DI)}{2} = \frac{\Pr_1(D)\Pr_1(I) + \Pr_2(D)\Pr_2(I)}{2},
\end{aligned}
$$

where the last identity uses the experts' judgements that D and I are independent. These independence judgements are preserved if and only if $\Pr(DI) = \Pr(D)\Pr(I)$, i.e. if and only if

$$
\frac{\Pr_1(D)\Pr_1(I) + \Pr_2(D)\Pr_2(I)}{2} = \frac{\Pr_1(D) + \Pr_2(D)}{2} \times \frac{\Pr_1(I) + \Pr_2(I)}{2}.
$$

By multiplying both sides of this equation by 4, developing the product on the right hand side, and simplifying, it follows that

$$\begin{aligned}
& \mathrm{Pr}_1(D)\,\mathrm{Pr}_1(I) + \mathrm{Pr}_2(D)\,\mathrm{Pr}_2(I) = \mathrm{Pr}_1(D)\,\mathrm{Pr}_2(I) + \mathrm{Pr}_2(D)\,\mathrm{Pr}_1(I) \\
\Leftrightarrow\ & \mathrm{Pr}_1(D)(\mathrm{Pr}_1(I) - \mathrm{Pr}_2(I)) = \mathrm{Pr}_2(D)(\mathrm{Pr}_1(I) - \mathrm{Pr}_2(I)) \\
\Leftrightarrow\ & (\mathrm{Pr}_1(D) - \mathrm{Pr}_2(D))(\mathrm{Pr}_1(I) - \mathrm{Pr}_2(I)) = 0.
\end{aligned}$$

The latter can hold only if $\mathrm{Pr}_1(D) = \mathrm{Pr}_2(D)$ or $\mathrm{Pr}_1(I) = \mathrm{Pr}_2(I)$, i.e. if the experts agree on the probability of drought or of political instability – which is not the case by assumption. So equal-weight linear pooling violates Independence Preservation. Similar violations can easily be constructed for non-equal weights (unless one individual i gets maximal weight $w_i = 1$).

While we have focussed on linear pooling as a way of aggregating probability judgements, the difficulty with preserving causal insights at the aggregate level is a very general one. Genest and Wagner [4] have shown that Independence Preservation is violated by many (linear or non-linear) probability aggregation rules, including geometric averaging, the most prominent alternative to linear averaging. Thus the difficulty of preserving causal knowledge is not an artifact of requiring Independence of Irrelevance Alternatives (a condition violated for instance by geometric averaging).

Genest and Wagner [4] regard this finding as evidence that Independence Preservation is not a reasonable condition. We would not like to go so far. In our view, those unanimous independence judgements that are grounded in unanimous causal judgements about the world should not be overruled. We take Genest and Wagner's impossibility finding not as a reason to abandon the goal of preserving judgements of independence, but as a reason to move to a two-stage approach that explicitly takes causal judgements into account. Causal judgements are aggregated first, and probabilistic ones only subsequently. Moreover the latter are aggregated in a way that differs from standard probability aggregation, namely in a way that is constrained by the causal judgements formed at the first stage. Our two-stage approach will satisfy a version of Independence Preservation restricted to unanimously held *causal* independencies.

3 Aggregating causal judgements

As before, let $\mathbf{V} = \{V, W, ...\}$ be a non-empty set of variables. In our example of the aid agency above, $\mathbf{V}$ contains the variables R (rainfall), Y (crop yields), P (political conflict) and F (famine). How can we represent judgements on how the variables in $\mathbf{V}$ are causally interrelated? Let us introduce a binary predicate symbol C to represent a causal relevance relation on $\mathbf{V}$, where, for any two variables V and W in $\mathbf{V}$, we write VCW to mean that V is causally relevant to W. In the case of the aid agency, an expert who thinks that rainfall is causally relevant to crop yield whereas political conflict isn't

would hold that RCY but not that PCY. A causal relevance relation C is called *acyclic* if, for any finite sequence $V_1, V_2, \ldots, V_k$ of variables in $\mathbf{V}$, it is not the case that

$$V_1 C V_2, V_2 C V_3, \ldots, V_{k-1} C V_k \text{ and } V_k C V_1.^8$$

A causal relevance relation C induces a *directed graph*, whose vertices are the variables in $\mathbf{V}$ and where, for any two variables V, W in $\mathbf{V}$, there is an edge from V in the direction of W if and only if VCW. This graph is a *directed acyclic graph* (DAG) if C is an acyclic relation.

A *Bayesian network* is a DAG with associated conditional probabilities: each variable in the graph is endowed with a conditional probability distribution given its parents in the graph. In this section, however, we set this quantitative information aside and focus on qualitative features of the DAG alone. In particular, we investigate how a group of individuals can arrive at an aggregate judgement on what the causal relevance relation C between the variables in $\mathbf{V}$ is.

Consider a group of n individuals, labelled $1, 2, ..., n$, each of whom holds a particular judgement on the nature of the causal relevance relation between the variables in $\mathbf{V}$. We write C_i to denote the causal relevance relation according to individual i's judgement. A combination of causal relevance relations across the n indivdiuals is called a *profile* and denoted $(C_1, C_2, ..., C_n)$. A *causal judgement aggregation rule* is a function that assigns to each profile $(C_1, C_2, ..., C_n)$ (in some domain of admissible profiles) a single aggregate causal relevance relation C.

To give some examples of causal judgement aggregation rules, consider the class of threshold rules. A *threshold rule*, with threshold k (where $1 \leq k \leq n$), assigns to each profile $(C_1, C_2, ..., C_n)$ the causal relevance relation C defined as follows: for any two variables V and W in $\mathbf{V}$,

$$VCW \Leftrightarrow \text{at least } k \text{ individuals have } VC_iW.$$

Examples of threshold rules are the *majority rule* ($k = \frac{n+1}{2}$), the *union rule* ($k = 1$) and the *intersection (or unanimity) rule* ($k = n$).

Are these satisfactory causal judgement aggregation rules? It is easy to see that each of these three rules has a considerable defect. The majority and union rules fail to ensure acyclicity of the aggregate causal relevance relation, even when all individuals hold acyclic such relations. To see this, suppose the aid agency consults three experts, with the following individual judgements. They all agree that rainfall is causally relevant to crop yields, but they disagree on the causal relations between the other variables. The first expert thinks that crop yields are causally relevant to famine, which is causally relevant to political conflict; the second thinks that famine is causally relevant to political conflict,

---

[8]Note that our definition of acyclicity also rules out cycles of length $k = 1$, i.e., we cannot have VCV for any variable V.

which is causally relevant to crop yields; and the third thinks that political conflict is causally relevant to crop yields, which is causally relevant to famine. In consequence, the causal relevance relation generated by the majority rule violates acyclicity: the relation contains a cycle from crop yields to famines to political conflict to crop yields. It is obvious that the union rule has the same defect. The intersection (or unanimity) rule, by contrast, ensures acyclicity of the aggregate causal relevance relation, but may generate a sparse or even empty such relation, with few variables deemed causally relevant to any others, whenever there are disagreements between the experts.

Although threshold rules are particularly salient examples of causal judgement aggregation rules, they are by no means the only ones. So let us once again adopt an axiomatic approach and look for rules satisfying certain conditions.

UD (*Universal Domain*) The causal judgement aggregation rule accepts as admissible any logically possible profile of acyclic causal relevance relations.

AC (*Acyclicity*) The aggregate causal relevance relation is always acyclic.

UB (*Unbiasedness*) For any two variables V and W in $\mathbf{V}$, the aggregate judgement on whether V is causally relevant to W depends only on individual judgements on whether V is causally relevant to W, and the aggregation rule is neutral between whether or not this is the case.[9]

ND (*Non-Dictatorship*) There does not exist a fixed individual such that, for every admissible profile of causal relevance relations, the aggregate causal relevance relation is the one held by that individual.

Although these conditions may seem natural at first sight, they are mutually inconsistent.

Theorem 1 *If* $\mathbf{V}$ *contains three or more variables, there exists no causal judgement aggregation rule satisfying UD, AC, UB and ND.*

We derive this result from a recent impossibility theorem by Dietrich and List [2] in the emerging literature on judgement aggregation (e.g., List and Pettit [8], Pauly and van Hees [10], Dietrich [1]).

What are the possible escape routes from this impossibility result?

---

[9]Formally, for any V and W in $\mathbf{V}$ and any admissible profiles $(C_1, C_2, ..., C_n)$ and $(C_1^*, C_2^*, ..., C_n^*)$, if [for all i, VC_iW if and only if not VC_i^*W] then [VCW if and only if not VC^*W]. This formal statement is slightly weaker than the informal one in the main text but implies the informal one together with UD and AC.

The first route: relaxing universal domain. We may use a causal judgement aggregation rule which accepts as admissible input not all logically possible profiles of acyclic causal relevance relations, but only those profiles that meet an additional structural condition – namely profiles that, informally speaking, reflect a certain amount of cohesion – if not consensus – across different individuals' causal judgements. The additional structural condition on profiles might be such that the majority rule, or perhaps some other threshold rule, never generates an aggregate causal relevance relation violating acyclicity; in this case the majority rule or threshold rule in question could be employed on this restricted domain of admissible profiles. We consider two structural conditions of this kind.

Temporal order restriction. Suppose the individuals agree on the temporal order in which the events captured by the variables in $\mathbf{V}$ occur. Suppose further they agree that a variable V can be causally relevant to another variable W only if V strictly precedes W in this temporal order. Call any profile of causal relevance relations that is consistent with *some* such agreement *temporal-order restricted.* Formally, a profile is *temporal-order restricted* if there exists *some* weak ordering of the variables in $\mathbf{V}$ (reflexive, transitive and connected binary relation on $\mathbf{V}$) such that, for every pair of variables V and W in $\mathbf{V}$, if some individual holds V to be causally relevant to W (i.e., some i holds $V C_i W$) then V strictly precedes W in that ordering. For any such profile, the causal relevance relation generated by any threshold rule is acyclic, no matter how low or high the threshold is. The temporal constraint on what causal relevance judgements are deemed admissible guarantees the absence of any causal cycles at both the individual and aggregate levels.

Unidimensional alignment. Another structural condition on profiles that ensures acyclical causal judgements at the aggregate level – here under the majority rule (or any threshold rule with a higher threshold) – is *unidimensional alignment* (List [6]). A profile of causal relevance relations is called *unidimensionally aligned* if the individuals can be linearly ordered from left to right such that, for every pair of variables V and W in $\mathbf{V}$, the individuals who hold that V is causally relevant to W (i.e., the individuals i with $V C_i W$) lie either all to the left or all to the right of those who hold that V is not causally relevant to W (i.e., the individuals i who do not have $V C_i W$). For any unidimensionally aligned profile, the causal relevance relation generated by the majority rule is acyclic and coincides with the causal relevance relation held by the median individual with respect to the left-right ordering of the individuals (or, if the number of individuals is even, it coincides with the intersection of the causal relevance relations held by the two median individuals).

It is an empirical matter whether a group of experts – either before or after suitable deliberation – exhibits sufficient agreement in their causal judgements to meet the condition of temporal order restriction or that of unidimensional alignment. The kind of temporal agreement required for temporal order restriction seems empirically plausible at least in some situations.

The second route: relaxing acyclicity. A logically possible way to avoid the impossibility result of Theorem 1 is to give up the requirement that the aggregate causal relevance relation be acyclic. This, however, would constitute a major departure from the consensus on the nature of causal relations, which are widely held to be acyclic (Pearl [11]).

The third route: relaxing unbiasedness. We may use a causal judgement aggregation rule which violates the condition of unbiasedness. Such a violation may take one of the following two forms.

A neutrality violation. There exist pairs of variables V and W in $\mathbf{V}$ such that the aggregation rule is not neutral between whether or not V is causally relevant to W. Examples of causal judgement aggregation rules violating neutrality are threshold rules with any threshold k different from simple majority. It can be shown that a threshold rule is guaranteed to generate an acyclic causal relevance relation if and only if the threshold k exceeds $\frac{m-1}{m}$, where m is the number of variables in $\mathbf{V}$. If the set of variables $\mathbf{V}$ is infinite, only the intersection (or unanimity) rule guarantees acyclicity at the aggregate level, but if $\mathbf{V}$ is finite, then a supermajority rule with a suitably high threshold is sufficient. A problem with this approach, as noted above, is that it may lead to sparse or even empty aggregate causal relevance relations unless the disagreement between experts is limited.

An independence violation. There exist pairs of variables V and W in $\mathbf{V}$ such that the aggregate judgement on whether V is causally relevant to W depends not only on individual judgements on whether V is causally relevant to W but also on individual judgements involving other variables. Examples of causal judgement aggregation rules violating independence are *sequential priority rules* (List [7]). Here the different possible pairs of variables are considered one-by-one in a given order (which may be chosen, for example, by some criterion of epistemic priority). On each pair of variables V, W, the aggregate judgement is then determined as follows. If the question of whether V is causally relevant to W is constrained by the aggregate judgements on pairs considered earlier (together with acyclicity), then the aggregate judgement on this question is made in accordance with those earlier constraints. But if it is not constrained in such a way, then the aggregate judgement on V's causal relevance to W is made by applying some voting method, such as majority voting, to the individual judgements on V vis-à-vis W. This approach guarantees acyclicity of the aggregate causal relevance relation, but at the expense of "path-dependence": the order in which causal judgements are made on different pairs of variables may determine what the causal relevance relation will look like. An agenda setter on a committee of experts may strategically exploit this feature of the causal judgement aggregation rule by proposing an order of priority among different pairs of variables that is likely to give rise to aggregate causal judgements that the agenda setter wants the committee to make.

The fourth route: relaxing non-dictatorship. A final way to avoid the impossibility result of Theorem 1 is to allow the aggregate causal relevance relation to be determined by an antecedently fixed individual: a "dictator". But since we are interested in making use of the information contained in the causal judgements of more than one individual, this is hardly an attractive solution to our aggregation problem. In some cases, however, may think that it is an acceptable compromise to appoint a trusted expert as the "dictator" for arriving at qualitative causal judgements – in the form of a DAG – while continuing "democratically" when it comes to determining the associated quantitative probability information at the second stage of our two-stage appraoch.

Which of the different possible escape routes from the impossibility result of Theorem 1 is compelling depends on details of the decision problem at hand, the nature of the disagreements between the experts, the level of trust we place in them, and so on. In the next section, we assume that through one of these routes – excluding that of relaxing acyclicity – a "consensus" on a causal relevance relation and thereby on a DAG has been achieved, and we turn to the question of how the associated conditional probabilities can be determined.

4   Aggregate probability judgements compatible with aggregate causal judgements

We have analysed how a group can arrive at an aggregate judgement on the causal relations between variables. We now assume that such an aggregate causal judgement has been reached through one of the routes just discussed and suppose that the group also seeks to make an aggregate probability judgement (about the variables taking various values), which is required to be compatible with that aggregate causal judgement.

In its most general form – ignoring for the moment the causal judgement – a probability judgement can be represented by a joint probability function over the variables in **V**. We assume that each variable can take finitely, or countably infinitely, many possible values. For example, we may distinguish between a particular number of possible levels of conflict. Let us label the variables $V_1, ..., V_m$. A joint probability function Pr assigns a probability $\Pr(v_1, ..., v_m) \geq 0$ to each combination $(v_1, ..., v_m)$ of values of these variables, where the sum of the probabilities is 1.

The joint probability $\Pr(v_1, ..., v_m)$ can be factorised into the product of conditional

probabilities:[10]

$$\begin{aligned}
\Pr(v_1, ..., v_m) &= \Pr(v_1)\Pr(v_2|v_1)\Pr(v_3|v_2, v_1)\cdots\Pr(v_m|v_{m-1}, ..., v_1) \\
&= \prod_{j=1}^{m}\Pr(v_j|v_1, ..., v_{j-1}). \qquad\qquad (1)
\end{aligned}$$

In our famine example, where V_1, V_2, V_3, V_4 are the levels of, respectively, rainfall, crop yield, political conflict and famine, we have

$$P(v_1, v_2, v_3, v_4) = P(v_1)P(v_2|v_1)P(v_3|v_1, v_2)P(v_4|v_1, v_2, v_3).$$

When is the probability judgement expressed by Pr *compatible* with a given causal judgement? Recall that a causal judgement takes the form of a particular directed acyclic graph (DAG) over the variables $V_1, ..., V_m$, with an arrow from V_j to V_k just in case V_j is considered causally relevant to V_k (V_jCV_k). For any variable V_j, we write $PA(V_j)$ to denote the list of V_j's parent variables in the graph, and we write $pa(V_j)$ to denote any list of values of these parent variables.[11]

For instance, suppose that, according to the consensus DAG in our famine example, no variable is causally relevant to rainfall (V_1); only rainfall (V_1) is causally relevant to crop yield (V_2); only crop yield (V_2) is causally relevant to political conflict (V_3); but both crop yield (V_2) and political conflict (V_3) are causally relevant to famine (V_4). Then $PA(V_1)$ contains no variable, $PA(V_2)$ contains precisely V_1, $PA(V_3)$ contains precisely V_2, and $PA(V_4)$ contains both V_2 and V_3.

Without loss of generality, suppose the variables $V_1, ..., V_m$ are labelled such that those with no parent come first, those with a parent but no grandparent come next, those with a grandparent but no great-grandparent come thereafter, and so on. If the original labelling $V_1, ..., V_m$ does not have this property, we can simply relabel the variables appropriately and replace the factorisation (1) by one using the new labelling. So the parents of any variable V_j come before V_j.[12] But of course not all of $V_1, ..., V_{j-1}$ need to be causally relevant to V_j. For instance, in our famine example V_2 but not V_1 is (directly) causally relevant to V_3. As causally irrelevant variables should have no effect on V_j, the conditional probability $\Pr(v_j|v_1, ..., v_{j-1})$ should be insensitive to the

---

[10]In this expression, the conditional probability $\Pr(v_j|v_1, ..., v_{j-1})$ can be derived from the joint probability function Pr via the formula $\Pr(v_j|v_1, ..., v_{j-1}) = \frac{\Pr(v_1, ..., v_j)}{\Pr(v_1, ..., v_{j-1})}$ (where $\Pr(v_1, ..., v_j)$ and $\Pr(v_1, ..., v_{j-1})$ are marginal probabilities derived from Pr), *provided* that $\Pr(v_1, ..., v_{j-1}) \neq 0$. If $\Pr(v_1, ..., v_{j-1}) = 0$, $\Pr(v_j|v_1, ..., v_{j-1})$ can be viewed either as undefined or as a primitive not derived from the function Pr. Under both interpretations, the factorisation (1) is still possible even if some $\Pr(v_1, ..., v_{j-1})$ is zero *whatever* value is substituted for $\Pr(v_j|v_1, ..., v_{j-1})$ (because some other factor on the right-hand side of (1) will be zero, as will be the left-hand side of (1)).

[11]So $pa(V_i)$ is any instantiation of $PA(V_j)$.

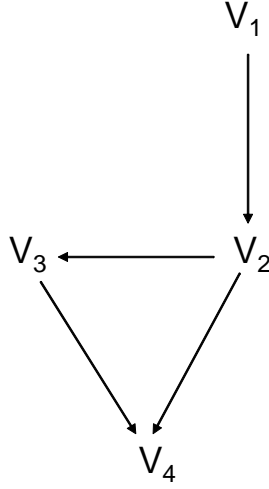[12]Formally, $PA(V_j)$ is a sublist of $(V_1, ..., V_{j-1})$.

Figure 1: Aggregate causal judgements in the famine example

non-parental values among $v_1, ..., v_{j-1}$. In other words, it should be sensitive only to the sublist $pa(V_j)$ of $v_1, ..., v_{j-1}$. Formally,

$$\Pr(v_j|v_1, ..., v_{j-1}) = \Pr(v_j|pa(V_j)).$$ (2)

We say that the probability judgement Pr is *compatible* with the given aggregate causal judgement if identity (2) holds for every variable V_j and every combination of values $v_1, ..., v_j$ with $\Pr(v_1, ..., v_{j-1}) \neq 0$.[13] The joint probability (1) then reduces to

$$\Pr(v_1, ..., v_m) = \prod_{j=1}^{m} \Pr(v_j|pa(V_j)).$$ (3)

For instance, in our famine example,

$$P(v_1, v_2, v_3, v_4) = P(v_1)P(v_2|v_1)P(v_3|v_2)P(v_4|v_2, v_3).$$

---

[13]There are different equivalent ways to define "compatibility" of Pr with the DAG. While we have defined compatibility by the *ordered Markov condition*, an equivalent definition uses the *parental Markov condition*: any variable is independent of its non-descendants given its parents. A third definition (chosen by Pearl) is by the validity of the decomposition (4). For the equivalence of these definitions, see Theorems 1.2.6 and 1.2.7 in Pearl [11].

5   A two-stage approach to the aggregation of probability judgements

As we seek to reach an aggregate probability judgement that is compatible with the aggregate causal judgement, the probability function Pr should satisfy the decomposition (3). The standard approach of directly aggregating individual probability functions

$$\Pr_1(v_1, ..., v_n), ..., \Pr_n(v_1, ..., v_n) \tag{4}$$

into an aggregate probability function $\Pr(v_1, ..., v_n)$ usually violates this decomposition. We propose to determine the aggregate probability function Pr by a two-stage approach. First, the individual causal judgements are merged into an aggregate causal judgement (the "consensus" DAG), as discussed above. Second, we consider the decomposition (3) relative to that aggregate causal judgement and determine each factor $\Pr(v_j | pa(V_j))$ through separate probability aggregation. The joint probability function $\Pr(v_1, ..., v_m)$ is then given by the product of these separately determined factors.

More formally, for *every* variable V_j in **V** and *every* combination $pa(V_j)$ of parental values, we merge the individual conditional probability functions

$$\Pr_1(v_j | pa(V_j)), .., \Pr_n(v_j | pa(V_j)) \tag{5}$$

into an aggregate conditional probability function $\Pr(v_j | pa(V_j))$. These separate aggregation exercises can each be performed, for example, by linear or geometric pooling. In our famine example, this involves

$$\begin{array}{lll}
 & \text{merging } \Pr_1(v_1), .., \Pr_n(v_1) & \text{into } \Pr(v_1), \\
\text{for any fixed } v_2, & \text{merging } \Pr_1(v_2 | v_1), .., \Pr_n(v_2 | v_1) & \text{into } \Pr(v_2 | v_1), \\
\text{for any fixed } v_3, & \text{merging } \Pr_1(v_3 | v_2), .., \Pr_n(v_3 | v_2) & \text{into } \Pr(v_3 | v_2), \\
\text{for any fixed } v_2, v_3, & \text{merging } \Pr_1(v_4 | v_2, v_3), .., \Pr_n(v_4 | v_2, v_3) & \text{into } \Pr(v_4 | v_2, v_3).
\end{array} \tag{6}$$

This two-stage approach has several distinctive properties, to which we now turn.

Compatibility with causal judgements.   The aggregate probability function Pr, given by (3), is automatically compatible with the aggregate causal judgement, represented by the appropriate DAG. In particular, Pr respects the causal Markov condition: any variable V_j is probabilistically independent of all its causal non-descendants given its causal parents (recall footnote 13). In our famine example, Pr makes political conflict independent of rainfall conditional on crop yield,[14] and famine independent of rainfall conditional on crop yield and political conflict.[15]   The causally motivated conditional

---

[14]Formally, $\Pr(v_1, v_3 | v_2) = \Pr(v_1 | v_2) \Pr(v_3 | v_2)$.

[15]Formally, $\Pr(v_1, v_4 | v_2, v_3) = \Pr(v_1 | v_2, v_3) \Pr(v_4 | v_2, v_3)$.

independencies are thus preserved, whereas other conditional independencies may or may not arise. By contrast, standard one-stage probability aggregation does not generally produce an aggregate probability judgement that respects any judgements of causal relevance.

Preservation of causal (conditional) independencies. What about the preservation of unanimously held independencies between variables (both conditional and unconditional ones)? Suppose, for example, that all individuals consider variables V_j and V_k probabilistically independent given V_l.[16] Does the aggregate probability judgement preserve this conditional independence? As we have seen, for standard probability pooling functions the answer is usually negative. Under our approach, by contrast, *causal* conditional independencies are preserved. To see why, suppose that all individuals judge V_j and V_k to be probabilistically independent given V_l because of a unanimous agreement that V_j's only causal parent is V_l and that V_k is not a causal descendant of V_j. Then the aggregate probability judgement respects this independence: according to Pr, V_j and V_k are also probabilistically independent given V_l.[17] The reason is that the first stage of our two-stage approach leads to an aggregate causal judgement that reflects the unanimous opinion on the causal relations between V_j, V_k, V_l,[18] and the second stage leads to a probability judgement that is compatible with this aggregate causal judgement.

Complexity reduction. Our two-stage approach subdivides an m-dimensional probability aggregation problem into several one-dimensional ones. Rather than aggregating joint probability functions over the vector $V_1, ..., V_m$ (of the form (4)), we aggregate conditional probability functions of a single variable V_j (of the form (5)). But we face several such aggregation problems, namely one for each variable V_j and each fixed combination of parent values $pa_j(V_j)$. This is less demanding on the side of individual inputs, as long as the aggregate DAG is not too rich in causal connections. To illustrate this complexity reduction, consider our famine example again, and suppose for simplicity that each variable can take only two values, i.e., there are only two levels of rainfall, two levels of crop yield, and so on. If we were to aggregate the joint probability functions $\Pr_i(v_1, v_2, v_3, v_4)$ directly, each individual would have to submit $2^4 - 1 = 15$ probability values (there are 2^4 possible combinations of values (v_1, v_2, v_3, v_4) but once the probabilities of $2^4 - 1$ of them are specified the remaining probability is given by one minus the sum of the rest). Specifying any one of these 15 probabilities is hard in practice: what, for example, is the probability of a combination of high rainfall *and* low crop yield *and* low political conflict *and* high famine? Under our approach, by contrast, each individual has to submit only probabilities of *single*-value events (possibly conditional such probabilities), like the

---

[16] Formally, $\Pr_i(v_j, v_k | v_l) = \Pr_i(v_j | v_l) \Pr_i(v_k | v_l)$.

[17] Formally, $\Pr(v_j, v_k | v_l) = \Pr(v_j | v_l) \Pr(v_k | v_l)$.

[18] Provided a "reasonable" causal judgement aggregation rule is used.

probability of high rainfall or the probability of high crop yield given low rainfall. The number of required probabilities is smaller than 15 in our example. Using (6), we can see that it equals

$$\sum_{j=1}^{4} \text{"number of possible values of } V_j \text{ minus 1"}$$
$$\times \text{"number of possible parent values } pa(V_j)\text{"}$$
$$= (2-1) \times 1 + (2-1) \times 2 + (2-1) \times 2 + (2-1) \times 2^2$$
$$= 1 + 2 + 2 + 4 = 9.$$

Type of informational input.   Our approach not only reduces the complexity of the aggregation problem; it also uses a different informational input from standard one-stage probability aggregation. First, we use the additional information of individuals' *causal* judgements – the information aggregated at stage one. Second, an interesting question arises about the nature of the probabilistic input used at stage two. Consider a variable V_j with parents $PA(V_j)$ in the aggregate DAG. As the aggregate DAG is the result of the aggregation of individual causal judgements, it is possible that some individuals do *not* agree that the variables listed in $PA(V_j)$ are the correct causal parents of V_j. They may think instead that not all of these variables are causally relevant to V_j or that some other variables are relevant although they are not included in $PA(V_j)$. In this case, how should we interpret the conditional probability $\Pr_i(v_j|pa(V_j))$ submitted at stage two by any such individual? In our famine example, what is an individual supposed to answer to the question "how probable is high political conflict given low crop yield?" if he or she actually thinks that famine rather than crop yield is causally relevant to political conflict? Two interpretations are possible, leading to different interpretations of our two-step approach. Either we interpret $\Pr_i(v_j|pa(V_j))$ simply as the standard conditional probability that individual i assigns to v_i given $pa(V_j)$, as derived from his or her own probability function $\Pr_i$. This conditional probability exists[19] even when $pa(V_j)$ does not contain the causally relevant values in i's judgement. Alternatively, we may interpret $\Pr_i(v_j|pa(V_j))$ as the probability that individual i *would* assign to v_j under the supposition that the variables in $PA(V_j)$ are the causally relevant ones (and take values $pa(V_j)$). Under this second interpretation, the individual entertains the hypothesis that the aggregate causal judgement is correct and expresses conditional probabilities based on this hypothesis. Here the probabilistic input used at the second stage of our two-stage approach differs substantially from that of standard probability aggregation.

---

[19]Provided that $\Pr_i(pa(V_j)) \neq 0$. In the special case of $\Pr_i(pa(V_j) = 0$, probabilities conditional on $pa(V_j)$ cannot be derived from $\Pr_i$ (see also footnote 10), and the further question arises as to what $\Pr_i(v_j|pa(V_j))$ then represents.

6 A challenge

The first stage of our two-stage approach restricts the second one by requiring the aggregate probability judgement to display certain conditional independencies reflecting the aggregate causal judgement. Roughly, the fewer causal links are accepted at the first stage, the more probabilistic independencies are enforced at the second stage. In the extreme case when *no* variable is deemed causally relevant to any other variable, the second stage produces an aggregate probability judgement according to which every variable is probabilistically independent of every other. Accepting few causal connections has the advantage of reducing the complexity of the probability aggregation problem at the second stage but the potential disadvantage of overrestricting the admissible probability assignments. This restriction is problematic when the sparse set of accepted causal links between variables is not a result of the individuals believing in sparse causal links but a result of a causal judgement aggregation rule setting a high threshold for the acceptance of causal links.

We are thus faced with a trade-off between the goal of reducing the complexity of the probability aggregation problem (achieved via a high threshold for accepting causal links between variables) and the goal of representing causal effects between variables whenever there are such effects (achieved via a low threshold for accepting causal links). We have argued that a high threshold for accepting causal links may help to prevent a cyclical aggregate causal judgement; whereas in other situations, particularly if the variables can be put into a temporal order, even a low threshold (perhaps lower than the majority threshold) guarantees acyclicity. We leave it as a challenge for future research to come up with causal judgement aggregation rules that perform well on both aspects of this trade-off: being neither too permissive nor too restrictive in accepting causal links while avoiding cyclical causal judgements.

References

[1] Dietrich, F. (2006) "Judgment aggregation: (im)possibility theorems", *Journal of Economic Theory* 126: 286-298

[2] Dietrich, F. and C. List (2006) "The impossibility of unbiased judgment aggregation", Working paper, London School of Economics

[3] Genest, C. and J. V. Zidek (1986) "Combining Probability Distributions: A Critique and Annotated Bibliography", *Statistical Science* 1 (1): 113-135

[4] Genest, C. and K. Wagner (1984) "Further Evidence against Independence Preservation in Expert Judgement Synthesis", *Technical Report* 84-10, Dept. of Statistics and Actuarial Science, University of Waterloo

[5] Lehrer, K. and C. Wagner (1981) *Rational Consensus in Science and Society*, Dordrecht: Reidel

[6] List, C. (2003) "A Possibility Theorem on Decisions over Multiple Propositions", *Mathematical Social Sciences* 45: 1-13; see also the corrigendum in *Mathematical Social Sciences* 52: 109-110 (2006)

[7] List, C. (2004) "A Model of Path-Dependence in Decisions over Multiple Propositions", *American Political Science Review* 98: 495-513

[8] List, C. and P. Pettit (2002) "Aggregating sets of judgments: an impossibility result", *Economics and Philosophy* 18: 89-110

[9] McConway (1981) "Marginalization and Linear Opinion Pools", *Journal of American Statistical Association* 76: 410-14

[10] Pauly, M. and M. van Hees (2006) "Logical constraints on judgment aggregation", *Journal of Philosophical Logic* 35: 569-585

[11] Pearl, J. (2000) *Causality: Models, Reasoning and Inference*, Cambridge: Cambridge University Press

[12] Spirtes, P., C. Glymour and R. Scheines (2000) *Causation, Prediction and Search*, 2nd ed., Cambridge MA: MIT Press

[13] Wagner, C. (1982) "Allocation, Lehrer Models, and the Consensus of Probabilities", *Theory and Decision* 14: 207-220

[14] Wagner, C. (1985) "On the Formal Properties of Weighted Averaging as a Method of Aggregation", *Synthese* 62: 97-108