A Formal Characterisation of Institutionalised Power

Andrew J.I. Jones [*] Marek Sergot [†]

June 1994; June 1995

Abstract

It is a standard feature of norm-governed institutions that designated agents are *empowered* to create particular kinds of states of affairs by means of the performance of specified types of actions. Frequently, the states of affairs are of a normative kind, in the sense that they pertain to rights and obligations, as for instance when a Head of Department signs a purchase agreement and thereby creates an obligation on his employer to pay for goods received. We use the term *institutionalised power* to stand for the notion of power we here seek to explicate. Following a lead from jurisprudential discussions of *legal power*, we distinguish *institutionalised power* from *permission* and *practical possibility*. We define a conditional connective intended to capture the consequence relation implicit in statements of the form: according to the constraints operative in institution s, the performance of some act A by agent x counts as a means of creating state of affairs B. When combined with deontic and action logics, the new connective facilitates the analysis of a number of notions crucial to the understanding of organised interaction in institutions, such as *authorisation* and *delegation*. We conclude with some illustrations of the expressive power of the new logical language.

1 Introduction

It is a commonplace feature of legal systems, and other norm-governed organisations, that particular agents are empowered to create certain types of states, by means of the performance of specified types of acts. Typically, the states created will have a normative character according to which obligations and rights are established for some agents vis-à-vis others, as for instance when a contract is made, or a marriage is effected, or ownership of an item is transferred. The performances by means of which these states are established will often be of a clearly prescribed, perhaps ritualised nature, involving the utterance of a particular form of words (e.g., the utterance of a specific type of performative

[*] Department of Philosophy and Norwegian Research Centre for Computers and Law, University of Oslo, Norway. e-mail: andrew.jones@filosofi.uio.no.

[†] Department of Computing, Imperial College of Science, Technology and Medicine, University of London, UK. e-mail: mjs@doc.ic.ac.uk.

sentence), or the production of a formal document, or the issuing of a pass, perhaps in a particular context (e.g., in the presence of witnesses). Within the jurisprudential literature there is a considerable amount of discussion of these matters, usually under the headings 'legal power', 'legal capacity' or 'norms of competence' (see [Lindahl 77, Ch. 6] and [Bulygin 92] for overviews); although what follows is intended in part to be a contribution to aspects of this jurisprudential issue, we need to make it explicit at the outset that 'empowering' is not an exclusively *legal* phenomenon, but is a standard feature of any norm-governed organisation where selected agents are assigned specific roles (in which they are empowered to conduct the business of that organisation). Thus although it is perhaps legal examples, such as making a contract, or decreeing a divorce, which come most immediately to mind, it is clear that illustrations of essentially the same sort of phenomenon also occur frequently in other than legal contexts: the Chief Librarian is empowered to waive lending restrictions; the Head of Department alone is empowered to assign duties to members of the department; only the student is empowered to register his chosen examination options; candidates for promotion can be put forward only by their section head. An adequate account of the mechanisms by means of which an organisation conducts its affairs will have to incorporate, one way or another, the phenomenon of *institutionalised power*, as we shall here choose to call it.

The problem of how to analyse institutionalised power arose for us in the contexts of legal knowledge representation and the formal specification of computer systems. It emerges, for instance, in our attempts to revise the Kanger-Lindahl framework for formalising Hohfeld's 'fundamental legal conceptions' (see [Jones & Sergot 92]). As we tried to show, that framework is a potentially very useful tool for giving precise specifications of agents' rights and obligations; to illustrate the point, we chose a domain in which a high level of precision is a requirement: the specification of *rights of access* to sensitive, confidential information. (The specific example concerned medical files in a psychiatric hospital [Ting 90] but any organisation in which there are rules defining security policies for various categories of personnel could have served our illustrative purposes just as well.) We showed there are good grounds for claiming that our techniques for mapping out classes of *normative positions*, as we call them, facilitate the identification of nuance and ambiguity in many rules about rights. However, the psychiatric hospital's security policies—in common, surely, with those operative in most complex organisations—also contained rules about which agents were empowered to assign rights, or to alter already existing rights. And it is difficult to see how the basic building blocks of the Kanger-Lindahl framework, which our approach retained, can be adequate to the task of representing the full meaning of such rules.

Those 'basic building blocks' are modal logics for the fundamental deontic notions of obligation and permission, and for the notion of successful action, according to which an agent brings it about that, or sees to it that, such-and-such is the case. (This approach to the logic of action is summarised in later sections.) It might be supposed that, of these notions, it is 'permission to do' which is closest to 'empowered to do'. But, as is pointed out by Makinson

[Makinson 86, p.408], jurisprudential theorists have long been aware that these two notions are not equivalent. Indeed, even if one were to enrich the logical language by adding a further modality, to express the idea of *agent ability*, or *practical possibility to act*, the notion of institutionalised power would still, it seems, remain uncaptured. Makinson cites a passage in which Hohfeld himself distinguished explicitly between:

(i) legal power,
(ii) the physical power needed to carry out the acts necessary for the exercise of a legal power, and
(iii) the permission to carry out those acts.

Makinson then adds a further illustrative example of his own:

> ...consider the case of a priest of a certain religion who does not have permission, according to instructions issued by the ecclesiastical authorities, to marry two people, only one of whom is of that religion, unless they both promise to bring up the children in that religion. He may nevertheless have the *power* to marry the couple even in the absence of such a promise, in the sense that if he goes ahead and performs the ceremony, it still counts as a valid act of marriage under the rules of the same church even though the priest may be subject to reprimand or more severe penalty for having performed it. [*Op. cit.*, p.409]

This is clearly a case in which the priest is empowered to marry the couple, but is not permitted to do so.[1]

Similarly, one may imagine circumstances in which it is not *practically possible* for the priest to marry the couple (because, say, he is sick or otherwise incapacitated), although he is still empowered to do so. Having the practical possibility to act is not a necessary condition for being empowered. Our position disagrees with that of Lindahl (see especially [Lindahl 77, pp206–210]), who suggested that the concept of practical possibility can be used in analysing legal power, provided that for cases of *legal* rather than mere *physical* power it is made explicit that the concern is with the practical possibility to create *normative* states. In our view, by contrast, an agent may have the practical possibility to create a normative state of affairs without being empowered. Consider here a point to which we shall return below, in our discussion of what we call *the transfer problem*: an agent may have the practical possibility to see to it that some couple are married (by, e.g., getting an authorised priest to conduct the ceremony) without himself being empowered to marry them. In short, the practical ability to perform acts which in one way or another create or lead to a normative

[1] *The Daily Telegraph*, 21/vi/93, contains a report about clandestine religious services conducted by former Roman Catholic priests who had left the priesthood to marry. According to the report, under canon law former priests "retain their sacramental powers but are forbidden to exercise them."

state of affairs is neither a necessary nor a sufficient condition for being empowered. Similar objections can be raised to a proposed formulation of (legal) power by Allen and Saxon (see, e.g., [Allen & Saxon 86, Allen & Saxon 93]).

Outside the domains of State or Church Law, one may easily find other examples illustrating the distinctions between the concepts in this triad. A departmental head may be empowered to purchase new equipment, but not be permitted to do so for equipment costing over \$50,000 unless he first consults the Finance Office. But his failure to consult them on a purchase of \$100,000 will not necessarily invalidate his purchase order. A member of academic staff may be empowered to sign and write "Accepted" on a student's supplementary reading list, thereby entitling the student to demand to be examined on the works on that list, but there may also be an obligation on the academic staff member not to exercise his power if the reading list is not properly representative of the subject area of the examination. But, again, even if the power is exercised in violation of the obligation, the staff member is empowered, and the student has his rights in regard to the examination. In both of the above examples, it is also clear that the agent concerned may be empowered and/or permitted without necessarily having the practical possibility to act.

There is reason to believe that the distinction between permission and institutionalised power is also important in understanding the notion of *authorisation*. Indeed it may be suggested that *authorisation* is ambiguous in regard to this pair. Sometimes when we say that an agent is authorised to do such-and-such we mean no more than that he has been granted permission to do it. But in other instances, as for example when we say that the Head of Department is authorised to purchase equipment, we mean first and foremost that he has been granted by the institution concerned the power to enter valid purchase agreements.

2 Preliminary analysis

The basic intuition which has guided our approach to the conceptualisation of institutionalised power is this: within institutions, organisations, or other normative systems, there operate constraints to the effect that the performance by some specified agent(s) x of some designated action is sufficient condition to guarantee that some specified agent y creates some (usually normative) state of affairs F. The agent y might be identical with the agent x, but this need not always be so. Often it would be appropriate to say that the agent y who creates the state of affairs F is the institution or normative system itself; for instance, it may be the registrar or priest who plays the role of x, performing the marriage ceremony, but it is the legal system or church which creates the normative relation of being married.

We are thus led to focus on statements of the following kind:

(i) "According to normative system/institution s,
if agent x sees to it that A then agent y sees to it that F."

We want to stress, however, that it is only a particular class of such statements that will correspond to the notion of institutionalised power.

Let the connective $\Rightarrow_s$ designate the notion of consequence we seek to capture. Using relativised modal sentences of the form $\mathrm{E}_x A$ to stand for "x sees to it that/brings it about that A", we reformulate statement (i) as

$$\text{(ii)} \qquad \mathrm{E}_x A \Rightarrow_s \mathrm{E}_y F$$

where x and y are assumed to be any of the agents who are designated by the institution s for the performance of the tasks specified (in the antecedent and consequent, respectively). The action modality E_x is discussed in the next section. Now, whatever other properties are exhibited by the logic of $\Rightarrow_s$, we should at least expect the following: given the truth of (ii) and that, on a given occasion, x does in fact see to it that A, then within the institution s the state of affairs described by F is established; in those circumstances, the situation described by F becomes established as a matter of institutional fact, relative to institution s. The idea here is, clearly, that (given (ii)) x's seeing to it that A must be an effective way of exercising power so that, relative to s, the state of affairs described by F becomes a matter of fact.

The relativisation to an institution of the new conditional connective is important for a number of reasons. What is taken as an established matter of fact (e.g., that agents a and b are divorced) relative to one institution s need not necessarily be so taken within the terms of reference of some other institution s'. Furthermore, as will be illustrated later in this paper, it is important for the purpose of characterising notions such as *delegation*, and some other organisational constructs, to allow for several different institutions in various forms of interaction.

The key question to be addressed concerns how the conditional connective $\Rightarrow_s$ is to be characterised. But before entering that discussion, consider whether the concern here is exclusively with those uses of $\Rightarrow_s$ which connect act-descriptions. Our answer is that we do not intend its use to be restricted in this way. We want to be able to express the idea that, within a given institution, certain acts *count as* means of creating certain kinds of normative states of affairs (perhaps 'institutionally significant states of affairs' might be a suitably neutral term); but we *also* want to be able to capture the fact that there are usually constraints within any institution according to which certain states of affairs of a given type *count as*, or *are to be classified as*, states of affairs of another given type. For example, it may be that an agent having such-and-such properties counts as an infant, or that a machine with such-and-such properties counts as a vehicle. Our conjecture is that the same conditional construction $\Rightarrow_s$ may be employed here as in the analysis of institutionalised power. Indeed it might be suggested that the *general* notion of institutionalised power concerns the contraints whereby an institution makes particular kinds of acts, or particular kinds of states of affairs, count as sufficient conditions for guaranteeing the applicability of particular *classificatory categories*; and these classifications, when made, often carry with them certain kinds of normative consequences concerning rights and duties.

How does this 'counts as' reading square with the account given in the opening remarks? The answer is perhaps easiest to see by reference to the special case of (ii)

(ii′) $$\mathrm{E}_x A \Rightarrow_s \mathrm{E}_x F$$

where the agent who brings about the designated state of affairs A is also the agent who establishes the state of affairs F. With the 'counts as' reading for the conditional $\Rightarrow_s$, statement (ii′) might be read as expressing that, relative to institution s, x's act of seeing to it that A counts as, or is to be regarded as, an act of establishing the state of affairs F, performed by x.

A similar reading for statements of the more general form (ii) is not quite so direct, but can be maintained nevertheless: x's act of bringing about A counts, within institution s, as a means by which agent y (who may but need not be the institution s itself) establishes state of affairs F; x may here be said to act on behalf of, or as an agent of, y—an interpretation which matches well the notion of 'being empowered' we seek to explicate. By way of illustration of a case where institution s and agents x and y are all distinct, consider a department s in which secretary x's signature counts as a signature of his or her boss y; here x is empowered in s to sign on behalf of y.

Put in the above terms, statement (ii′) is very reminiscent of Goldman's notion of *conventional generation*:

> Conventional generation is characterized by the existence of rules, conventions, or social practices in virtue of which an act A' can be ascribed to an agent [x, say], given his performance of another act, A. [Goldman 70, p25]

If we take it that Goldman's 'rules, conventions, or social practices' play the role of our 'institution' s then the similarity is striking. However, there are some clear differences between Goldman's 'conventional generation' and the notion of institutionalised power we seek to capture. Although we reserve more detailed comparison for future work, we note here the following two points:

(a) Goldman does not consider cases where the antecedent and the consequent of the generation connective are relativised to distinct agents. But for 'institutionalised power' it is important not to restrict our 'counts as' connective in this way.

(b) Goldman gives examples of conventional generation which would not appropriately be described as cases of *empowering*. For instance, x's act A of breaking a promise conventionally generates x's act A' of "doing what he ought not to do", but we would not want to say that x is here *empowered* to do what he ought not to do. Note further that where 'x does B' logically entails 'x does A' then we would presumably want to conclude that x's doing B *also* conventionally generates x's act A' of doing what he ought not to do. But we shall argue below that the notion of power is not transferable in this way across logical entailment.

3 On the action component E_x

As already indicated, we employ the same action modality E_x both for expressing that agent x creates/establishes states of affairs, and for expressing that x performs designated acts, rituals, and so forth. We have found reasons to be uneasy regarding this kind of dual employment of the action modality, and at times it may be that we are stretching the application of this approach to the logic of action to its limits. Nevertheless, we postpone discussion of alternative logics of action to future investigation. For present purposes we shall assume that the logic of E_x is that of a (relativised) classical modal system of the type Chellas [Chellas 80] calls the system E, i.e., the smallest system containing PL and closed under the rule RE:

RE. $\quad \dfrac{A \leftrightarrow B}{\mathrm{E}_x A \leftrightarrow \mathrm{E}_x B}$

with the additional axiom schema

T. $\quad \mathrm{E}_x A \rightarrow A$

and the additional rule of inference

R¬N. $\quad \dfrac{A}{\neg \mathrm{E}_x A}$

The schema T is naturally a feature of any logic of *successful* action. The rule R¬N captures part of the idea that agent x is in some sense responsible for bringing about the state of affairs A, that it might have been otherwise but for his actions. Whatever else we might have in mind (cf. the discussion in [Pörn 77, Elgesem 93]), on no account could we accept that an agent brings about what is logically true.

A good case can also be made (cf. [Elgesem 93]) for adopting the axiom schema

C. $\quad (\mathrm{E}_x A \wedge \mathrm{E}_x B) \rightarrow \mathrm{E}_x (A \wedge B)$

though not for its converse

M. $\quad \mathrm{E}_x (A \wedge B) \rightarrow (\mathrm{E}_x A \wedge \mathrm{E}_x B)$

since, with PL and closure under logical equivalence RE, M would yield

RM. $\quad \dfrac{A \rightarrow B}{\mathrm{E}_x A \rightarrow \mathrm{E}_x B}$

RM is unacceptable: since any sentence A logically implies $\top$ where $\top$ is a tautology, RM would yield the result that $\mathrm{E}_x A$ entails $\mathrm{E}_x \top$; but since we accept R¬N this is turn would give the result that $\neg \mathrm{E}_x A$.

We leave aside discussion of other possible properties of E_x , except to remark on the following two schemas, since they relate to later discussion on problems of transfer of power. We have as a special case of schema T, the schema:

$$\mathrm{E}_x \mathrm{E}_y A \rightarrow \mathrm{E}_y A$$

It is also possible, depending on the reading ascribed to E_x, to make a case for

$$\mathrm{E}_x \, \mathrm{E}_y \, A \rightarrow \mathrm{E}_x \, A$$

We do *not* assume validity of this schema in what follows, but neither do we rely on its absence. As a general principle we do not want our account of power to be dependent on detailed choices for the logic of the underlying action component. In fact, the only specific properties of E_x we make use of in what follows are validity of the schema T and closure under logical equivalence, RE.

4 On the logic of the connective $\Rightarrow_s$

Any normal conditional logic is, as Chellas puts it, "...normal with respect to the consequent" [Chellas 80, p269]; so were the logic of $\Rightarrow_s$ to be a normal conditional logic, that would mean (among other things) that the consequent is closed under logical consequence, i.e., that one accepts the rule:

$$\text{RCM.} \qquad \frac{B \rightarrow B'}{(A \Rightarrow_s B) \rightarrow (A \Rightarrow_s B')}$$

We view this rule as inappropriate for the notion of 'counts as' we seek to explicate. Suppose x's doing A counts (in s) as a means of securing that s sees to it that F; should it also then count as a means of securing the truth of any classical consequence of "s sees to it that F"? Surely not: imagine that x's uttering the words "I pronounce you man and wife" counts (in s) as a means of guaranteeing that s sees to it that a and b are married. It would then be bizarre to conclude that x's utterance act would also count in s as a means of guaranteeing that either Nixon is impeached or s sees to it that a and b are married. Furthermore, since any sentence B logically implies $\top$ where $\top$ is a tautology, acceptance of RCM would mean that if A counts in s as B, then A counts as any tautology.

Related to this is the further feature that mere logical consequence does not guarantee consequence of the 'counts as' type. That is, we reject for $\Rightarrow_s$ the rule:

$$\text{RI.} \qquad \frac{A \rightarrow B}{A \Rightarrow_s B}$$

(Note, however, that the rule RI is not a property of *all* normal conditional logics.)

However, the consequent of a $\Rightarrow_s$ conditional should certainly be required to be closed under logical equivalence, as should the antecedent. In which case we are led to accept, as a first shot, a conditional logic of type CE (in the Chellas classification, [Chellas 80, p270]), containing PL and the two rules:

$$\text{RCEC.} \qquad \frac{B \leftrightarrow B'}{(A \Rightarrow_s B) \leftrightarrow (A \Rightarrow_s B')}$$

$$\text{RCEA.} \qquad \frac{A \leftrightarrow A'}{(A \Rightarrow_s B) \leftrightarrow (A' \Rightarrow_s B)}$$

Can further rules or axiom schemas be adopted? Our inclination is to accept the schema:

CC. $((A \Rightarrow_s B) \wedge (A \Rightarrow_s C)) \rightarrow (A \Rightarrow_s (B \wedge C))$

and also to accept the schema:

CA. $((A \Rightarrow_s B) \wedge (C \Rightarrow_s B)) \rightarrow ((A \vee C) \Rightarrow_s B)$

Note that the converse of CC, that is, the schema

CM. $(A \Rightarrow_s (B \wedge C)) \rightarrow ((A \Rightarrow_s B) \wedge (A \Rightarrow_s C))$

could not be accepted: since it is a rule of PL that

$$\frac{B \rightarrow C}{B \leftrightarrow (B \wedge C)}$$

and since we accept RCEC, acceptance of CM would generate the closure property (of the $\Rightarrow_s$ consequent, under logical consequence), rule RCM, which we seek to avoid.

In a parallel fashion we shall argue against acceptance of the converse of CA; here, the argument relates to an important aspect of institutionalised power, to which we earlier alluded under the label 'the transfer problem.' The point is essentially this: suppose that x is empowered to marry couple a and b by performing ritual R. Now suppose that some other agent y brings it about that x performs ritual R—y, let us imagine, successfully exercises influence over x by some means or other. So x performs the ritual and the couple a and b are married. Despite his successful exercise of influence, we would not here want to say that y too was empowered, by institution s, to marry the couple. Institutionalised power is not *transferable* in that way.

It is worth noting, in passing, that our logical language is capable of expressing different notions of power; y, who let us say has effective influence over x (who is empowered), certainly has a *form* of power, because he can successfully control the behaviour of an institutionally empowered agent. And we might wish to go on, to distinguish those cases where the means employed in y's control over x is legal/illegal with respect to institution s (by which x is empowered). But our main point now is that the effective control (over x) that y has does not guarantee that y also inherits the institutionalised power which the institution has bestowed on x.

Returning to the logical issue regarding the converse of CA, the first point to note is that it is clearly a logical truth that if y brings it about that x performs ritual R, then x performs ritual R. More generally, since the T schema is an axiom of the action logic, then all instances of the following schema are logically true:

$$\mathrm{E}_y \, \mathrm{E}_x \, A \rightarrow \mathrm{E}_x \, A$$

Recognition of the transfer problem thus requires that the connective $\Rightarrow_s$ must not be constrained by the following rule:

PTR. $$\frac{A \rightarrow B}{(B \Rightarrow_s C) \rightarrow (A \Rightarrow_s C)}$$

Now PL contains the following rule:

$$\frac{A \rightarrow B}{B \leftrightarrow (A \vee B)}$$

which together with the converse of CA would generate rule PTR—assuming, as we do, RCEA.

From the semantical point of view, the above classical conditional logic may be modelled in terms of a minimal conditional model[2]

$$\mathcal{M} = \langle W, f_s, P \rangle$$

where W and P are understood as in standard models, as a set of possible worlds and as a valuation for each of the atomic sentences in these worlds, respectively, and f_s is a function which (relative to a given institution or normative system s) assigns a set of propositions to each proposition (subset of W) at each world. The basic understanding of sentences of the form $A \Rightarrow_s B$—the idea that within, or according to, institution s, the state of affairs described by A counts as one of the sort described by B—is captured by the truth condition (where $\alpha \in W$):

$$\models_{\alpha}^{\mathcal{M}} A \Rightarrow_s B \quad \text{iff} \quad \|B\|^{\mathcal{M}} \in f_s(\alpha, \|A\|^{\mathcal{M}})$$

So, for instance, if at α the institution s designates x's seeing to it that A as a means of guaranteeing that some agent y (perhaps x, perhaps the institution itself, perhaps some other agent) sees to it that B, then the proposition expressed by the sentence "y sees to it that B" will be among those picked out by f_s at α for the proposition expressed by "x sees to it that A".

To secure the validity (for any institution s) of the schemas CC and CA, we adopt (respectively) the following constraints on the function f_s (for all sets X, Y, Z, all $\alpha \in W$):

(cc) if $Y \in f_s(\alpha, X)$ and $Z \in f_s(\alpha, X)$ then $Y \cap Z \in f_s(\alpha, X)$
(ca) if $X \in f_s(\alpha, Y)$ and $X \in f_s(\alpha, Z)$ then $X \in f_s(\alpha, Y \cup Z)$

We leave open the possibility that there may be true instances of the schemas

I. $A \Rightarrow_s A$
Sym. $(A \Rightarrow_s B) \wedge (B \Rightarrow_s A)$

although they are of little import in regard to the present application; in the context of institutionalised power, most instances of I and Sym will no doubt ordinarily be false, but we shall not rule out the possibility of exceptions. These are further respects in which our analysis of $\Rightarrow_s$ differs from Goldman's account of conventional generation, referred to in section 2, since Goldman's relation is assumed to be both irreflexive and asymmetric.

[2]We have omitted here the components required for the E_x modality. They are added straightforwardly and independently of the other parts, as functions e_x from W to $\mathrm{Pow}(\mathrm{Pow}(W))$ with the appropriate model conditions to validate T and R$\neg$N. For the general strategy, see, e.g., [Chellas 80, Ch.7]. For a slightly different approach in the same spirit see [Elgesem 93].

Goldman's relation of conventional generation is also assumed to be transitive. The counterpart within the logic of $\Rightarrow_s$ would be the transitivity, or hypothetical syllogism, schema:

$$\text{S.} \qquad (A \Rightarrow_s B) \rightarrow ((B \Rightarrow_s C) \rightarrow (A \Rightarrow_s C))$$

Although we need to subject this principle to further investigation, we have been unable to produce any convincing counter-instances and are inclined to accept it. Consider again an example mentioned above: suppose secretary x's signature counts in s as boss y's signature, and suppose further that boss y's signature counts in s as a means of producing a valid claim for expenses. Then surely secretary x's signature counts in s as a means of producing a valid claim for expenses. Alleged counter-instances, we surmise, will turn out on closer examination to concern situations in which one or both of the premisses is/are false, rather than casting doubt on the validity of the transitivity principle itself.

Schema S may be added to the principles already adopted for the logic of $\Rightarrow_s$ without undesirable consequences. The obvious model condition to adopt in order to validate the schema S is the following:

$$\text{(s)} \qquad \text{if } \; Y \in f_s(\alpha, X) \; \text{ and } \; Z \in f_s(\alpha, Y) \; \text{ then } \; Z \in f_s(\alpha, X)$$

Even if it were to transpire that convincing counter-examples to S could be found, a weakened form of transitivity:

$$\text{SD.} \qquad (A \Rightarrow_s B) \rightarrow ((B \Rightarrow_s C) \rightarrow \mathrm{D}_s(A \rightarrow C))$$

will nevertheless be a truth of the logic. Here D_s designates a more general kind of institutional constraint, to be discussed in the next section.

5 Institutional constraints in general

The connective $\Rightarrow_s$ has been proposed as a means of representing conditionality of the 'counts as' kind. However, it is clear that not all conditional sentences true of a given institution s will be of that sort. Among the others will surely be conditionals which describe relations of logical consequence, of causal consequence, and of deontic consequence. To illustrate these three in turn, consider first the logical relation mentioned above:

$$\mathrm{E}_y\,\mathrm{E}_x\,A \rightarrow \mathrm{E}_x\,A$$

which is clearly a constraint on any institution s, in the sense that it is incompatible with what s does that $\mathrm{E}_y\,\mathrm{E}_x\,A$ and $\neg\mathrm{E}_x\,A$. Second, any given institution s will be subject to causal constraints; for instance, it may be true of s (but not necessarily of other institutions) that if a given agent x sees to it that A then this causes some other agent y to see to it that B. Thirdly, a given institution s will be subject to deontic constraints: it may, for instance, be true of s (but not necessarily of other institutions) that if x sees to it that A then y ought to see to it that C.

Although we shall not here enter into the details of how *these* types of conditionals are to be represented, our suggestion will be that they share with 'counts as' conditionals the property of being *constraints on an institution*; given the truth of a conditional sentence "if A then B" of any of these types, we shall maintain that a true relativised necessity statement of the form

$$\mathrm{D}_s(A \rightarrow B)$$

may be derived, where s is the institution under consideration and, as before, $\rightarrow$ is the material conditional. Thus, for the specific case of the 'counts as' conditional, this relationship to the *general* notion of institutional constraint is secured by adopting the axiom schema:

$\Rightarrow_s$D. $\quad (A \Rightarrow_s B) \rightarrow \mathrm{D}_s(A \rightarrow B)$

We read expressions of the form $\mathrm{D}_s(A \rightarrow B)$ as "it is a constraint of (operative in) institution s that if A then B", or as "it is incompatible with the constraints operative in s that A and not-B."

The next question is of course how the D_s modality is to be defined. Our provisional proposal is to assign D_s the properties of a (relativised) normal modality of type KD.

Two immediate consequences of the choice of a normal modality are of course that, for any s, all logical truths will be constraints of s, and, for any s, the logical consequences of that which constrains s also constrain s. The logic contains the K-schema:

DK. $\quad \mathrm{D}_s(A \rightarrow B) \rightarrow (\mathrm{D}_s A \rightarrow \mathrm{D}_s B)$

The adoption of the D-schema:

DD. $\quad \mathrm{D}_s A \rightarrow \neg \mathrm{D}_s \neg A$

or (equivalently, because D_s is normal) the schema

DP. $\quad \neg \mathrm{D}_s \bot$

imposes a consistency requirement on the constraints operative in any s. For reasons that have already been indicated, the stronger T-schema:

DT. $\quad \mathrm{D}_s A \rightarrow A$

cannot be adopted. In particular, we want to allow the consistent assertion of

$$\mathrm{D}_s(A \rightarrow B) \wedge \mathrm{D}_{s'}(A \rightarrow \neg B)$$

for $s \neq s'$, even where A is in fact true. Distinct institutions may exhibit, or be subject to, mutually incompatible constraints.

Semantically, the introduction of D_s is straightforward, either by adding components of a standard model or by using a minimal model of the appropriate sort. In terms of standard model components for D_s, the model is (with E_x components again omitted for simplicity):

$$\mathcal{M} = \langle W, f_s, d_s, P \rangle$$

where W, P and f_s are as before, and where d_s is a function from W to subsets of W: $d_s(\alpha)$ is the set of worlds in W which are D_S-accessible from α. The relevant truth condition is thus (for all $\alpha \in W$):

$$\models_\alpha^{\mathcal{M}} \mathrm{D}_S\, A \quad \text{iff} \quad d_s(\alpha) \subseteq \|A\|^{\mathcal{M}}$$

The model conditions that validate DD and $\Rightarrow_s$D are, respectively:

(dd) $\quad d_s(\alpha) \neq \emptyset$

(sd) $\quad$ if $Y \in f_s(\alpha, X)$ then $d_s(\alpha) \cap X \subseteq Y$

It might be considered undesirable to mix minimal and standard models in this fashion; a minimal model of the following form can be used instead:

$$\mathcal{M} = \langle W, f_s, f_s^D, P \rangle$$

f_s^D is a function from W to Pow(Pow(W)): for every $\alpha \in W$, $f_s^D(\alpha)$ is the set of propositions (subsets of W) that are D_S-necessary at α. The corresponding truth condition is then (where $\alpha \in W$):

$$\models_\alpha^{\mathcal{M}} \mathrm{D}_S\, A \quad \text{iff} \quad \|A\|^{\mathcal{M}} \in f_s^D(\alpha)$$

For D_S to be normal this model needs to be 'augmented' as regards f_s^D [Chellas 80, Ch.7], i.e., it should satisfy the following two conditions, for all X, Y, α:

(dm) $\quad$ if $X \subseteq Y$ and $X \in f_s^D(\alpha)$ then $Y \in f_s^D(\alpha)$

(dci) $\quad \bigcap f_s^D(\alpha) \in f_s^D(\alpha)$

or, equivalently, the condition

(da) $\quad X \in f_s^D(\alpha)$ iff $\bigcap f_s^D(\alpha) \subseteq X$

For a model of this type, the conditions needed to validate DD, DP and $\Rightarrow_s$D are, respectively (for all X, Y, and all $\alpha \in W$):

(dd′) $\quad$ if $X \in f_s^D(\alpha)$ then $-X \notin f_s^D(\alpha)$

(dp′) $\quad \emptyset \notin f_s^D(\alpha)$

(sd′) $\quad$ if $Y \in f_s(\alpha, X)$ then $(-X \cup Y) \in f_s^D(\alpha)$

Above (section 2) we specified a form of 'detachment' requirement on the logic of the 'counts as' conditional to the effect that, given the truth of $A \Rightarrow_s B$ and the truth of A, it must follow that, *within the institution* s, the state of affairs described by B becomes established as a matter of fact. We are now in a position to propose a means of securing a 'detachment' result of this sort. From $\Rightarrow_s$D and DK it obviously follows that

$\Rightarrow_s$DK. $\quad (A \Rightarrow_s B) \to (\mathrm{D}_S\, A \to \mathrm{D}_S\, B)$

While this schema is itself too weak to capture the detachment property we seek, it nevertheless provides a step in the right direction. To see why, suppose again that there is a 'counts as' relation in s of the form $\mathrm{E}_x\, A \Rightarrow_s \mathrm{E}_y\, B$. Then it may be claimed that x's seeing to it that A is an act of special significance in s; its performance must be taken to constrain the properties of s: just because the 'counts as' conditional is true, x's seeing to it that A assumes the status of an event which constrains s. Thus, given $\mathrm{E}_x A \Rightarrow_s \mathrm{E}_y B$ and $\mathrm{E}_x A$ we should be able to draw the conclusion that $\mathrm{D}_s \mathrm{E}_x A$.

Generalising this proposal, we adopt the axiom schema:

$$\text{Const.} \qquad (A \Rightarrow_s B) \rightarrow (A \rightarrow \mathrm{D}_s A)$$

which together with $\Rightarrow_s$DK yields

$$\text{Det} \Rightarrow_s . \qquad (A \Rightarrow_s B) \rightarrow (A \rightarrow \mathrm{D}_s B)$$

Since D_s is normal, we may now derive, for instance,

$$\text{DEx.} \qquad (\mathrm{E}_x A \Rightarrow_s \mathrm{E}_y F) \rightarrow (\mathrm{E}_x A \rightarrow \mathrm{D}_s F)$$

and our detachment requirement is met.[3]

In connection with our proposed solution to the detachment problem it is important to note that there is no sentence B such that, for all sentences A, $A \Rightarrow_s B$ is a theorem. Thus the adoption of the axiom schema Const does not yield the intuitively unacceptable result that $A \rightarrow \mathrm{D}_s A$ becomes a theorem of the logic.

It is possible to envisage other properties of D_s and further strengthening of the link between 'counts as' and institutional constraints in general. For example, both of the following axiom schemas seem plausible:

$$\mathrm{D}_s A \rightarrow \mathrm{D}_s \mathrm{D}_s A$$
$$(A \Rightarrow_s B) \rightarrow \mathrm{D}_s (A \Rightarrow_s B)$$

We have still to examine the effect of such refinements.

6 Examples and further developments

As indicated in our opening remarks, the notion of institutionalised power is normally encountered in combination with other concepts, notably the fundamental deontic concepts of obligation and permission, denoted here by the operator O and its dual P respectively, and practical possibility (to act), designated here by Can .

These concepts can be inserted more or less straightforwardly into the existing framework since they are largely independent of the building blocks we have employed to analyse institutionalised power: deontic logic has of course been studied extensively (which is not to say that all problems have been resolved); proposals for the treatment of practical possibility are also available (see, e.g., [Pörn 89]). We omit the details here. Our purpose in this section is just to

[3]It has been pointed out to us that 'counts as' consequences follow non-defeasibly on this analysis. We have chosen for present purposes to put to one side the (clearly important) issue of defeasibility which we do not see as essential to the analysis of the 'counts as' relation itself.

give some indication of the expressive richness of the multi-modal language that results from such extensions. Examples are given for illustrative purposes only: we make no attempt here to provide a systematic presentation of the points arising.

It can be argued that a formal language for representing institutionalised power must provide some form of quantification, at least over agents, because of the nature of what is being represented. We have put this refinement to one side in order to focus on the notion of power itself. If the reader finds this completely unacceptable, no harm is done if examples in this section are read in such a way that the agent index stands for a *class* of agents (priests, heads of department, faculty members, managers) rather than individuals.

Consider first the priest p who is empowered to marry some couple by performing ritual r. Let m stand for "the couple are married" and suppose the priest's power is represented thus:

$$(\mathrm{m}_1) \qquad \mathrm{E}_p\, r \Rightarrow_s \mathrm{E}_s\, m$$

The case where a (former) priest p is forbidden to exercise his power to create a (valid) marriage can be represented as follows:

$$(\mathrm{m}_2) \qquad (\mathrm{E}_p\, r \Rightarrow_s \mathrm{E}_s\, m) \wedge \neg \mathrm{P}\, \mathrm{E}_p\, r$$

Note that this last expression does not imply $\neg \mathrm{P}\, \mathrm{E}_p\, m$: the state of affairs represented by

$$(\mathrm{m}_3) \qquad (\mathrm{E}_p\, r \Rightarrow_s \mathrm{E}_s\, m) \wedge \neg \mathrm{P}\, \mathrm{E}_p\, r \;\wedge\; \mathrm{P}\, \mathrm{E}_p\, m$$

is consistent. Here, the priest p is forbidden by the Church to exercise his own power to create the marriage, but the Church does not go so far as to prohibit all action which p might take to bring about the marriage; p is not forbidden, in particular, to act as a layman in this matter, and to see to it that some other empowered agent performs the marriage.

In similar style, the case where an empowered priest p is permitted to exercise his power, but is not practically able to do so, is represented:

$$(\mathrm{m}_4) \qquad (\mathrm{E}_p\, r \Rightarrow_s \mathrm{E}_s\, m) \wedge \mathrm{P}\, \mathrm{E}_p\, r \;\wedge\; \neg \mathrm{Can}\, \mathrm{E}_p\, r$$

Note again that this expression does not imply $\neg \mathrm{Can}\, \mathrm{E}_p\, m$, even if there is no other designated ritual for creating the marriage besides bringing it about that r: it may still be practically possible for priest p to influence (bring about) the exercise of some other agent's power to create the marriage.

We turn now to some examples of different forms of 'effective power'. Consider the administrative arrangement found in many University departments whereby the power to submit a research grant proposal resides with the Head of Department (h, say). A grant proposal is produced (g) when the Head h adds his signature to the completed proposal form (represented by his bringing about state of affairs f):

$$\mathrm{E}_h\, f \Rightarrow_s \mathrm{E}_s\, g$$

Various forms of effective power possessed by some other member of the faculty, a, can be distinguished. Compare the following:

$$(p_1) \quad \text{Can}\, E_a\, E_h\, f \qquad \text{(whether or not also } P\, E_a\, E_h\, f\text{)}$$

$$(p_2) \quad \text{Can}\, E_a\, O\, E_h\, f$$

$$(p_3) \quad E_a\, d \Rightarrow_{s'} E_{s'}\, O\, E_h\, f$$

(where $E_a\, d$ is "a gives h a completed proposal form for signature"). In the last case, the institution s' corresponds (say) to 'the department', which need not be the same institution as that, s, which acts on the Head of Department's power to submit proposals. From the 'exercise of power' schema DEx, (p_3) yields:

$$(p_4) \quad E_a\, d \rightarrow D_{s'}\, O\, E_h\, f$$

which says that if faculty member a acts accordingly, then, from the viewpoint of the department s', the Head of Department h is placed under an obligation to exercise his power to create a valid proposal.

It seems to us that this is a realistic representation of a common organisational arrangement. Grant proposals might require the signature of the Head of Department before they will be accepted, but refusal by the Head of Department to add his signature to an otherwise acceptable proposal would ordinarily be regarded—within the department—as an abuse of the power the Head of Department wields in virtue of his position. It might also be that the institution s which empowers the Head of Department to submit grant proposals also recognises the Head of Department's position vis-à-vis the other faculty members in these matters; in that case, the institutionalised constraint

$$(p_5) \quad E_a\, d \Rightarrow_s E_s\, O\, E_h\, f$$

would also obtain.

We close with a brief illustration of one common 'authorisation' scenario. Imagine that employee a is empowered, by making the appropriate demand, to make a valid claim for refund of his travelling expenses, on condition that he is in possession of an authorising note from his boss b; only boss b is empowered to issue valid notes (by adding his signature, let us say). Aspects of this scenario might be represented by two constraints, thus:

$$(a_1) \quad D_s(n \rightarrow (E_a\, d \Rightarrow_s E_s\, c))$$

$$(a_2) \quad E_b\, i \Rightarrow_s E_s\, n$$

Here n stands for "employee a possesses a valid note", d stands for "a demand for expenses is made", c stands for "a valid claim is made", and i stands for "the boss adds his signature".

Does it follow that, in virtue of these arrangements, boss b is *empowered* to empower employee a to make a valid claim? To answer the question, note that, although we can deduce the following from (a_1)–(a_2):

$$(a_3) \quad D_s(E_b\, i \rightarrow (E_a\, d \Rightarrow_s E_s\, c))$$

we cannot derive that boss b is here empowered; for *that* would require a constraint of the form

$$(a_4) \qquad \mathrm{E}_b\, i \Rightarrow_s \mathrm{E}_s\,(\mathrm{E}_a\, d \Rightarrow_s \mathrm{E}_s\, c)$$

It is possible that (a_4) does hold, i.e., that the boss is empowered in this way, but it does not follow from the scenario as formulated: (a_4) implies (a_3) but is not implied by (a_1)–(a_2).

Of course, the *effect* of the boss's signing the note, $\mathrm{E}_b\, i$, is the same whether the authorisation structure is that represented by (a_1)–(a_2) or (a_4); in the former case there is a conditional constraint in the institution s empowering the boss to 'trigger' or 'enable' a's power to claim—but this is not yet a constraint of the 'counts as' type that characterises instances of institutionalised power.

The last example touches on questions concerning the treatment of *conditional* power: it is clear that in general it is the performance of a prescribed act or ritual *in specified circumstances* C which guarantees that state of affairs F is established. There are several possibilities for expressing such conditional structures with the resources available (one of which is illustrated above). Exploration of these possibilities, and identification of possible shortcomings and further extensions, are the main focus of our current investigations.

Acknowledgements

Andrew Jones' work has been carried out within the ESPRIT Basic Research Project No. 6471, MEDLAR II, *Mechanising Deduction in the Logics of Practical Reasoning* and Marek Sergot's within ESRC/MRC/SERC Joint Council Initiative project G9212036. The work was also supported by the Portuguese research programme STRIDE, FEDER and JNICT under the project JURAD, and by the ESPRIT Working Group 8319 MODELAGE. We thank members of these various project groups, Layman Allen, and Ingmar Pörn for their helpful comments on earlier drafts of this paper.

References

[Allen & Saxon 86] L.E. Allen and C.S. Saxon. Analysis of the logical structure of legal rules by a modernized and formalized version of Hohfeld fundamental legal conceptions. In A.A Martino and F. Socci, editors, *Automated Analysis of Legal Texts*, pages 385–451. North-Holland, Amsterdam, 1986.

[Allen & Saxon 93] L.E. Allen and C.S. Saxon. A-Hohfeld: A language for robust structural representation of knowledge in the legal domain to build interpretation-assistance expert systems. Ch. 8 in *Deontic Logic in Computer Science: Normative System Specification*, J.-J.Ch. Meyer and R.J. Wieringa (eds). John Wiley & Sons, 1993.

[Bulygin 92] E. Bulygin. On norms of competence. *Law and Philosophy*, 11:201–216, 1992.

[Chellas 80] B.F. Chellas. *Modal Logic—An Introduction.* Cambridge University Press, Cambridge, 1980.

[Elgesem 93] D. Elgesem. *Action Theory and Modal Logic.* Doctoral Thesis, University of Oslo, Department of Philosophy, 1993.

[Goldman 70] A.I. Goldman. *A Theory of Human Action.* Prentice-Hall, New Jersey, 1970.

[Jones & Sergot 92] A.J.I. Jones and M.J. Sergot. Formal specification of security requirements using the theory of normative positions. In Y. Deswarte, G. Eizenberg and J.-J. Quisquater, editors, *Computer Security—ESORICS 92*, Lecture Notes in Computer Science 648, pages 103–121. Springer-Verlag, Berlin Heidelberg, 1992.

[Lindahl 77] L. Lindahl. *Position and Change—A Study in Law and Logic.* Synthese Library 112, D.Reidel, Dordrecht, 1977.

[Makinson 86] D. Makinson. On the formal representation of rights relations. *Journal of Philosophical Logic*, 15:403–425, 1986.

[Pörn 70] I. Pörn. *The Logic of Power.* Blackwells, Oxford, 1970.

[Pörn 77] I. Pörn. *Action Theory and Social Science: Some Formal Models.* Synthese Library 120, D. Reidel, Dordrecht, 1977.

[Pörn 89] I. Pörn. On the nature of a social order. In J.E. Fenstad et al., editors, *Logic, Methodology and Philosophy of Science VIII.* Elsevier Science Publishers, 1989, 553-567.

[Ting 90] T.C. Ting. Application Information Security Semantics: A Case of Mental Health Delivery. In D.L. Spooner and C.E. Landwehr, editors, *Database Security: Status and Prospects III.* North Holland, Amsterdam, 1990.