

The Structure and Evaluation of Planning Arguments

THORBJOERN MANN

903 Washington Street
Tallahassee, Florida 32303
USA
thormann@nettally.com

Abstract: The structure of ‘planning arguments’—arguments commonly used in discussion about plans and policy proposals—is discussed. Based on the conceptual framework of the ‘argumentative model of planning’ proposed by H. Rittel, an approach for their systematic and transparent evaluation by discourse participants is presented. Procedural implications for its application in the planning process are discussed, and the potential for information technology support for such processes explored.

Résumé:

Keywords: Argument, discourse, evaluation, planning, plausibility, policy.

1. Rationale for the development of an argument assessment technique

The problem addressed in this article is the systematic and transparent evaluation of arguments used in discussions about design and planning proposals and in general in policy discussion of all kinds. The exploration of these ideas was triggered by the “design methods” teachings of Professor H. Rittel at the University of California at Berkeley and his ideas of the “argumentative model” as the basis for information systems for design and planning, followed by the development of “issue based information systems” (IBIS) at the Studiengruppe für Systemforschung in Heidelberg. This work was motivated in part by the desire to help design and planning professions to “catch up” with developments in other disciplines, mainly with respect to improved processes, information support, and the use of the computer (which was at the time only in its infancy and only available to larger institutions and firms). Computer

modeling and the development of computerized data banks were quickly becoming the domain of a new breed of experts, pointing to examples such as the successes of the space effort.

In urban planning, however, precisely this expert model had produced a number of controversial projects that proved disastrous, failing to address user needs and concerns. These failures contributed to the growing demands for greater citizen participation, even “user design,” but at the very least for more transparency in how planning decisions are made: the expectation that decisions be based on due consideration of user and citizen concerns and arguments. The slogan to base decisions on arguments was supported by political opposition movements as well as by theory—for example, J. Habermas’s Communicative Action ideas (Habermas 1971, 1973, 1984).

For design, and especially for the effort to provide better information support for design and planning, this resulted in a few specific challenges:

- In addition to the usual expert and textbook knowledge and statistical data, the information systems would have to accommodate the information distributed in the population potentially affected by proposed plans.
- The information should not only support the decision-making process in a more direct manner than traditional “reference” information systems, but also do this in such a way that the basis for decisions would be more “transparent,” that is, more explicitly linked to the reasons or pros and cons for and against plan proposals contributed by participants. This meant, among other things, that these information systems would have to be much more “process”-oriented, and based on a realistic model of the activity they were supporting.

Rittel’s answer to these challenges was the “argumentative model” of design and planning, on which he based his concept for “issue based information systems”—IBIS—or “argumentative planning information systems”—APIS.¹ This model, which aligned itself with popular demands for significant social projects and poli-

¹The seminal article for this was Kunz and Rittel (1970), followed by a number of prototype studies carried out under Rittel’s guidance at the Studiengruppe für Systemforschung in Heidelberg and Berkeley (Rittel 1980). The concepts have since become the basis for a variety of planning support programs, for example by “gIBIS” at Conklin’s CogNexus firm, the “Zeno Argumentation Framework” (Gordon 1999), or the programs “bCisive” and “Rationale” offered by the Australian group Aus-think, to mention but a few.

cies to be decided on the merit of issues and arguments, carried the implied expectation of some systematic and transparent process for the evaluation of the arguments brought forward in the course of such project discussions. However, the techniques available for such careful evaluation fell far short of meeting this expectation. While elaborate techniques for the formal evaluation of competing plan proposals had been developed, such as the evaluation procedure for buildings by Musso and Rittel (1969), or the widely adopted Cost-Benefit-Analysis, these techniques were disassociated from the familiar traditional process of parliamentary procedure, the model for even “countercultural” events such as town hall meetings in which distributed information from the citizenry could be brought to bear on projects. Apart from other criticisms of CBA (such as the glossing over potential disagreements among affected parties about the goals that define the benefits, or the problems associated with the assessment of intangible values such as impacts on people’s health or lives) these formal evaluation tools share the feature that their application requires trained experts, and that they are carried out in a different conceptual frame of reference from that of the discussion of pros and cons, especially in public settings such as town hall meetings. It would require affected and interested parties from the non-expert public to learn a specialized vocabulary in order to participate in the application of such formal evaluation tools.

In more “participatory” meetings, however, the tradition was restricted to the rules of debate with, inevitably, subsequent decision-making by voting. The problem with voting is obviously that the process itself does not provide any traceable connection between the arguments and the vote. The lack of this connection all too often led to significant distortion of the basic idea of parliamentary process, in that votes can be taken along predetermined party lines, in blatant disregard of arguments, for reasons of political power relations, or—if one would prefer a more benevolent interpretation—because decision-makers do not have any good tools for a more systematic and transparent evaluation of arguments. For this reason, the assumption and aim of the investigation presented here was to look for a way in which the arguments brought into a discussion or debate could be evaluated in a manner that more directly would indicate their support for the position that would eventually be decided upon. Another way of describing this concern might be: to attempt to refine the familiar opinion polling practice to the point where it addresses the merit of arguments supporting or attacking a proposal, not just expressing overall support or rejection. The assessment of the merit of design and planning arguments became the main topic for my work in this area (Mann 1977), which was initially conceived as just a minor contribution to the argumentative model project. Over time, however, I came to see it as a much more important general concern far beyond the realm of

much more important general concern far beyond the realm of architecture and planning.

Proposed Assessment Approach for Planning Arguments

One might have expected that formal or informal logic, rhetoric, or “critical thinking” texts would have produced sufficient resources and building blocks for the construction of such argument assessment tools. However, the majority of the work in these disciplines studied at the time seemed curiously unhelpful for this project. For one, such texts mainly focused on the analysis and evaluation of individual arguments, not on the multiple “pros and cons” one finds in planning discussions. The preoccupation was with the criterion of “truth” of premises and conclusion. There are reasons for suspecting that “truth” does not apply very well to planning arguments that contain deontic (“ought”) claims as premises. The truth of a factual claim is very different from the acceptability, appropriateness or desirability of a plan proposal and the goals and concerns it addresses, which, after all does not satisfy one vital criterion for truth—that of its claim having a correspondence in the real world. A proposed plan is debated precisely because it is not true yet. It also turned out that the typical planning argument does not meet formal logic standards for validity—it is at best “inconclusive.” This may have been the reason that I could not find much evidence of efforts devoted to the analysis of the particular argument patterns used in design and planning, and to the development of better evaluation techniques for the way all the pros and cons support or fail to support the overall decision about a proposal. A main difference between the studies found in the literature and the perceived challenge of the study was that analyses in logic and even deontic logic tried to establish general standards for what argument forms should be considered “valid” in general, by anybody (and therefore should be used in argumentation) and which forms lacked merit (and therefore should be avoided or disregarded). In contrast, the problem for planning was to find ways for discourse participants to express and communicate about the merit that each participant attributes to an argument.

The approach I developed, briefly, consisted of the identification and description of the structure of the typical planning argument and the proposal to have decision-makers or evaluators judge all premises according to their (apparent, to the evaluator) plausibility rather than truth. I then suggested that the plausibility of an argument could be expressed as a function of the degrees of plausibility of its premises and that of its structure or reasoning rule. The weight of each argument is then seen as a function of the argument’s assigned plausibility and the weight of relative importance of the goals or objectives referenced in the deontic premise of the

argument. Finally, the plausibility of the plan proposal was expressed as a function of the argument weights of all the arguments offered in the discussion. The adoption of such a tool in the planning process would require some procedural changes, which will be outlined, and the question is discussed how such a modified procedure might be supported by technology.

It will be necessary to provide some background of the argumentative model as proposed and its vocabulary, as proposed by H. Rittel, to facilitate the further discussion.

2. The Argumentative model of planning

Questions and Issues

Rittel (1977, 1980) suggested that design and planning can be seen as a process in which questions are raised—either explicitly in a discussion among several participants, or internally by the designer or planner working on a problem. Answers to these questions are searched for, proposed, and discussed. Issues are understood as questions that have become controversial in that different participants adopt incompatible positions as their answers. They then seek to support these positions by means of arguments and further answers (evidence), and to weaken opposing answers and arguments with counter-arguments.

Issue types

Different types of questions and issues are distinguished, corresponding to different types of claims that play a role in the discussion.

Most important in the design and planning discourse is the deontic claim and deontic issue. (D-claim and or D-issue). While it can be expressed in many different ways, it can usually be condensed or translated into a form such as “X ought to be” (claim) or “Should X be implemented?” (issue). The plan proposal is a deontic claim, as is (at least) one of the premises of such arguments. Deontic claims are usually discussed in the literature in terms of “goals” or “objectives,” “needs,” “norms” or “requirements.”

The proposal claim and the deontic claim of a design argument are linked by a factual or factual-instrumental claim (FI-claim), which, if challenged, can give rise to a corresponding FI-question or issue. A typical example is “(It is a fact that) X will cause Y” or “X will be instrumental to achieve Y”.

Sometimes such claims are qualified by adding specific conditions under which the instrumental claim holds: “X will cause Y given conditions C”. The claim about conditions C then calls for an

associated (often implicit) factual claim (F-claim) that these conditions are given.

Further question types in the planning discourse are the straight Instrumental questions (I-questions): “How can X be achieved?” and Explanatory questions (E-questions) that can include such questions as “What is the cause of (some condition) z?” “What is z?” or “What is the definition (or meaning) of z?”

Positions

The positions adopted by discourse participants about controversial questions are usually “Yes” and “No”—affirmative or negative answers to both deontic and factual issues. Significant is the third position “Inadequate question” or “Wrong question” which suggests that the question as posed is based on an inadequate or inappropriate understanding of the problem or controversy at hand; often accompanied by a suggestion of the better question or problem view: “The real problem is” Rittel emphasized that many successful resolutions of controversies resulted from the adoption of such a new or different problem view—one on which the opposing parties could reach agreement.

Argument merit as the basis for decisions

Underlying the argumentative model is the assumption that much of the pertinent information needed for planning is distributed in the population—the various parties affected by the problem—rather than only as textbook and expert knowledge or data banks. Such information must be obtained through discourse—involving participation by affected parties in the planning process—and specifically through argument. The corresponding expectation is that the decisions about the plan should then be based on the merit of those arguments. This raises the question of how one might or should go about assessing argument merit, and how the assessment can be made explicit and transparent, in the interest of accountability of decision-makers who make such judgments and decisions on behalf of many other people. It is this question I sought to explore as a contribution to the overall project of the argumentative model.

3. The Design and Planning Argument: Components and Structure

The call for a systematic and transparent argument assessment procedure for design and planning arguments must begin with consideration of the components and structure of these arguments, especially since they turn out to be slightly different from the typical

arguments one usually finds analyzed in logic, rhetoric, and critical thinking texts.

The conclusion or plan proposal

The ‘conclusion’ of the planning argument is, of course, the deontic claim (D-claim) that the plan proposal be adopted, or that some detail be adopted in the plan description. In the following the proposed plan X will be denoted as ‘D(X)’—for example: “Plan X ought to be implemented”.

The premises

In the usual discussion of planning proposals, as in much common conversation, arguments are often expressed informally in incomplete manner: many premises that the speaker assumes can be “taken for granted” are left out, and often other information is added that consists of further evidence or support for one of the key premises of the argument—one which also may be left unspoken even though it is further supported. A systematic argument evaluation must, however, explicitly consider all the key premises, even those that are only implied. But the assessment of an individual argument should address only the main premises: inasmuch as it also may contain additional support or evidence for one or more of the premises, the claim of the respective premise constitutes a separate “successor” issue that may have to be discussed and clarified (if challenged) before returning to the main issue. This means that in this view, some of the argument components suggested by Toulmin (1958) in his argument diagram—e.g. backing, qualifier—are not part of the core planning argument but of successor issues. The premises of a complete basic planning argument are the following:

- The Factual-Instrumental premise suggesting a link or relationship REL between the proposed plan X and some consequences or implication Y of the implementation of X, (a goal or objective suggested as desired by the argument proponent);
- FI(X REL Y) or FI{(X REL Y)|C}—the claim that the relationship holds given certain conditions C;
- The Deontic premise D(Y)—claiming that the consequence or goal Y should be pursued;
- The Factual premise F(C)—claiming that the conditions C are indeed present.

The argument structure or reasoning rule

The structure of the typical complete planning argument—in other words, its inference or reasoning rule R, is the following:

D(X) (Conclusion, plan proposal)
 because
 FI(X REL Y) (Factual-Instrumental premise)
 and
 D(Y) (Deontic premise)

Respectively, for arguments specifying the conditions C under which the FI-premise is assumed to hold:

D(X)
 because
 FI{(X REL Y)|C}
 and
 D(Y)
 and
 F(C)

This inference rule (which some writers, for example Galle (1996, 1997) have labeled “abductive,” an assessment I do not share since in my understanding the abductive reasoning pattern involves factual, not deontic claims) does not meet formal logic standards for validity, in the sense of a classical valid syllogism that guarantees the truth of the conclusion if all premises are true. Others (for example Zenker (2009)) label this ‘pro and con’ argumentation “conductive,” acknowledging the multitude of pros and cons “leading” to the conclusion. For the time being at least, I would like to stay with the label “standard design argument” or “standard planning argument,” for these patterns involving deontic claims to support deontic decisions—if only to distinguish them from similar arguments involving only factual claims and conclusions. The pattern is, at best, “inconclusive” from a formal logic point of view. This defect does not prevent anybody from using such arguments, however, nor from according it considerable merit.

The inference rule as stated is its general form, implying assertion of all claims. Of course, each claim and claim component can be modified as negation. The various possible combinations of assertion and negations in the rule give rise to a great many variations of the pattern, most of which are no more conclusive than the above all-assertion form, but some will be outright self-contradictory. In addition, there are many relationship types that can be used in the rule: causation, analogy, similarity, identity, equivalence, even just “reminiscent of,” or whether a feature “fits” a person’s or community’s character or style, are just a few examples. The variety of relationship types can generate thousands of variations of the rule. It will therefore be necessary to always include the assessment of the inference rule as used in a particular

argument, in its evaluation.

The planning argument inference rule shares the inconclusiveness property with the inductive reasoning pattern. For induction—which plays such an important role in scientific reasoning—probability theory and statistical inference calculating the probability of the “null hypothesis” (the assertion that a hypothesis suspecting a causal relationship between two variables, for example, is *not* true and observed evidence only due to random variation) have overcome the logical shortcomings of the induction inference rule. Unlike induction, the planning argument pattern has not been the subject of extensive discussion and analysis in the textbooks. Most do not even acknowledge it.² Beyond its inconclusive character, this may have something to do with the deontic claims it contains, that cannot easily be labeled “true” or “false” in the same way as the factual claims, and with the aspect that planning decisions never rest on only one such argument but on many pros and cons that must be considered together. But in spite of the ubiquitous assurances by decision-makers to “carefully weigh the pros and cons,” these decision-makers would be hard put to explain just how they would do this.

It may be necessary to emphasize again that the proposed approach does not aim at offering recommendations for what kinds of

² This impression is admittedly not based on any exhaustive scrutiny of all literature in logic, rhetoric, or informal logic, but on the survey of some main texts and courses in “critical thinking,” for example the well-respected self-study courses by Hall or Zarefsky offered by The Teaching Company, where one would have expected such developments to be at least mentioned. The predominant concern still appears to be with recommending what arguments people *should* consider valid and plausible, as opposed to the task of having people express (for further discussion) the *actual* assessments that determine their decisions, whether or not these assessments may be sufficiently supported, rational or deliberated, in the view of outside argumentation experts. Efforts to develop generally valid reasoning rules are of course important for the development of “expert systems” aiming at substituting machine judgment for human judgment. But such aims are very different from that of devising procedures to support discourse to inform human judgment. In such procedures, the rules laid out by theory would either have to be known by all participants (as a result of general education) or themselves become the substance of arguments that raise “successor issues” within the debate—about the plausibility of the inference rule participants should assign in the evaluation. While I cannot argue with the goal of teaching better reasoning skills in general education, I also believe a significant part of the inference rule assessment will be about its applicability to the specific premises in the particular situation, and thus a matter for discussion within that situation.

arguments *should* be offered by participants in planning discourse, but rather to simply facilitate and improve how arguments actually brought forward in such discussions may be more carefully evaluated and transparently connected with the eventual decisions (for example, “votes”) by individual participants. This may explain why there is no treatment or mention of such topics as “stock issues,” a staple concern of rhetoric. Of course, arguments about stock issues must be evaluated once they are offered in a discussion. But the recommendation to explore and introduce stock issue arguments in a debate is part of a more general rhetoric or argumentation theory, not of that of the evaluation of any such arguments.

4. The assessment of planning arguments

The task of argument assessment is approached from the assumption that the merit of an argument must be derived from the merit of its components, the premises and inference rule. The first problem here is that of the appropriate criteria to apply to this task.

Criteria for argument assessment

Truth

Traditional logic is concerned with the truth of premises, as the prime guarantee that the conclusion will be true. Following Aristotle, there can be only two values for this assessment: “true” and “false.” It was already suggested above that this criterion does not really apply well to the deontic claims in the planning argument, unless we are ready to accept different kinds of truth: e.g. factual and “moral” truths. In which case the question arises as to whether the “moral” truth of deontic premises guarantees the truth of the conclusion in the same way as factual truth. There is also the added difficulty that many factual claims about the future (which are present in all arguments about plans to be adopted and implemented in the future) do not allow people to assert complete certainty about them: there are degrees of conviction involved that make the binary assessment values “true/false” inappropriate. This points out an important aspect of such assessments: the decisions we want to derive from them do not rest on the actual truth or falsity of the state of affairs to which the claims refer. We do not have complete and certain knowledge about them—not even about factual claims regarding laws of nature (which often are invoked in the factual-instrumental premises of planning arguments). As Popper (1968) has argued, we cannot conclusively prove a scientific hypothesis, only disprove it, and if all our best attempts at refuting it fail, accept the hypothesis “provisionally” with some degree of confi-

dence, until a better hypothesis is proposed. This means that when we make decisions based on such assessments, they are based on our subjective degree of belief or confidence in their truth. It suggests that we express our assessment on a scale that should allow conveying anything between complete agreement or certainty about a claim, and complete disagreement. Any form of expressing our assessments of argument premises should properly be understood as expressions of that subjective degree of confidence.

Probability

In science, the need for judgments with degrees of certainty is met with the concept of probability, expressed on a scale ranging from zero to one (or zero to a hundred “percent”). For the factual-instrumental premise, an appropriate criterion will be an estimate of the probability that it will hold for the implementation and consequences of the proposed plan. Probability estimates would also apply to factual premises for the future. Probability of evidence being observed under the hypothesis or its opposite null-hypothesis (the hypothesis that observed evidence is just a result of random deviations), has been extensively studied as the basis for inferential statistics in scientific hypothesis-testing. But speaking about the probability of deontic claims (goals, for example, or the desirability of certain consequences of our planned activities) would be awkward and inappropriate; it conflates the problem of estimating or predicting what decision-makers *will* do given certain information, with the issue of what they *ought* to do.

Plausibility

A “weaker” term that would be applicable to not only deontic claims, but to all the types of claims we encounter in the planning argument is ‘plausibility’. I propose to use it in a way that encompasses both truth and probability, and acknowledges a deontic claim as an understandable and acceptable (however debatable) one. The proposed approach will use this term for expressing assessments about all those types of premises: as expressions of participants’ personal, subjective degree of belief in the acceptability of any of these claims as part of the construction of reasons to justify or reject action, in spite of lack of complete certainty. In this interpretation, ‘plausibility’ should not be understood as “quasi-factual.” But it can even be used to evaluate the appropriateness and applicability of the reasoning rule to the case at hand—judgments that commonly are couched in terms of its “validity.” It may be argued that two separate judgments should be used for this latter purpose: One for the plausibility (in the sense of ‘validity’) of the inference rule itself, and one for its applicability or fit to the situation at hand. But since validity is also a binary judgment (a

sylllogism is either valid or not) and already has been ruled out as applying to this inference pattern, it seems acceptable to conflate the two judgments of plausibility and applicability into one “plausibility” score. Again, it must be emphasized that such assessment of the argument pattern is necessary because, among the many variations of the rule, there are some that are shakier than others, and some entirely self-contradictory; and their applicability to the specific argument in the case at hand is not a matter of course.

Scale

The next question concerns the scale on which such judgments would be expressed. It might be best—to fully convey the difference between such subjective assessments from any kind of objective measurement—to only use qualitative statements such as “completely plausible,” “somewhat plausible” and “entirely implausible,” and it should again be made very clear that any quantitative-looking scale that might be adopted for practical purposes really does not “mean” anything more than subjective qualitative assessments. Such a scale, furthermore, should have a “midpoint” expressing the judgment of “don’t know” or “can’t make an assessment.” This is very different from the midpoint of 50% on the probability scale, for example, which expresses considerable confidence about the probability of an event such as a coin toss, rather than ignorance about it. A plausibility scale of pl ranging from -1 (complete implausibility) to $+1$ (complete plausibility) with the midpoint of zero (“don’t know”) is suggested as a first approximation for communicating about our plausibility judgments on argument claims.

Significance, weight, importance

Two arguments whose plausibility is considered equal by some participant, may nevertheless carry different *weight* or significance in that person’s mind, and therefore influence the person’s overall decision differently. This difference has its root in the different degree of importance associated (again, in the mind of the particular individual) with the deontic claims of the two arguments. The deontic premises can be seen as concerns or goals (of achieving or avoiding certain consequences of the planned action X), and this can be expressed in a ranking, or better, a weight of relative importance, of those goals—specifically, of all the deontics that have been mentioned in all the pro and con arguments about the proposed plan. Using a technique such as the Churchman-Ackoff Approximate Measure of Value (Ackoff, Churchman and Smith, 1954), each participant can establish a set of weights w for these deontic claims and express those on a scale of zero to one such that

$0 \leq w(i) \leq 1.0$ and $\sum w(i) = 1.0$. In a group, these can become the subject of a separate discussion in the spirit of the Delphi Method, but all participants will use their own weights in the subsequent assessment.

Argument plausibility

On the basis of the assumption that argument plausibility will depend on the participants' judgments of plausibility assessments of its premises, it is now possible to express this dependency of argument plausibility ARGPL as

$$\text{ARGPL}(i) = f \{ \text{pl}(\text{FI}(\text{X REL Y})), \text{pl}(\text{DY}), \text{pl}(\text{R}) \} \quad \{1\}$$

or: the plausibility of an argument i is a function of the plausibilities of its factual-instrumental premise, the plausibility of its deontic premise, and the plausibility of the inference rule R as applicable to the argument. Two tentative, potential functions for expressing this relationship are suggested. The first is Argument plausibility function type I:

$$\text{ARGPL}(i) = \text{pl}(\text{FI}(\text{X REL Y})) * \text{pl}(\text{DY}) * \text{pl}(\text{R}) \quad \{2\}$$

and, respectively,

$$\text{ARGPL}(i) = \text{pl}(\text{FI}(\text{X REL Y})|C)) * \text{pl}(\text{DY}) * \text{pl}(\text{F}(C)) * \text{pl}(\text{R}) \quad \{3\}$$

for the expanded argument specifying condition C under which the FI relation REL holds, and the factual claim about that condition. These functions will result in argument plausibility scores that tend the more towards zero (undecided), the lower or closer to zero are the individual premise plausibility assignments. While this feature may disappoint expectations for such assessments to result in clear and decisive overall scores, it does realistically stress the importance of having adequate support for each premise to guarantee decisively strong assessment.

Equations 2 and 3 as stated apply only to the condition that all plausibility scores are on the positive side of the +1 / -1 scale, or that only one is negative, which turns the result into a negative score. As soon as two or more pl-scores are negative, additional assessments are needed to determine whether the evaluating person will consider that argument as a "pro" or a "con" argument (both interpretations are possible in some instances) or dismiss it as meaningless with an overall plausibility of zero. The interesting aspect of this question is the possibility that an argument can fail to become a significant component of the overall evaluation not because of its factual content but because of its structure. The infor-

mation presented in the premises might be quite significant and useful when arranged in a different argument pattern, or, for example, related to different context conditions under which the claimed instrumental relationship holds. This aspect undoubtedly deserves more discussion.

A more severe function (in that it is based on the assessment of only *one* of the premises), expressing an attitude such as “an argument is only as strong as its weakest premise,” is Argument function type II:

$$\text{ARGPL}(i) = \text{MIN} \{ \text{pl}(\text{FI}(\text{XRELY})), \text{pl}(\text{DY}), \text{pl}(\text{R}) \} \quad \{4\}$$

and, respectively,

$$\text{ARGPL}(i) = \text{MIN} \{ \text{pl}(\text{FI}(\text{X REL Y})|\text{C}), \text{pl}(\text{DY}), \text{pl}(\text{F}(\text{C})), \text{pl}(\text{R}) \} \quad \{5\}$$

Argument weight

The weight of each argument $\text{AW}(i)$ can now be considered a function of argument plausibility and the weight $w(\text{Di})$ of its deontic premise $\text{D}(i)$, generally:

$$\text{AW}(i) = f \{ \text{ARGPL}(i), w(\text{Di}) \} \quad \{6\}$$

for example:

$$\text{AW}(i) = \text{ARGPL}(i) * w(\text{Di}) \quad \{7\}$$

The form of this equation {7} might lead to misunderstanding of the concept of argument weight as a form of Expected Value, and to suggestions for applying Bayesian analysis or Subjective Expected Utility models to its calculation. Inasmuch as Expected Value involves probability estimates, this would apply only to the factual premises in planning arguments, that is, to the plausibility assessment of those premises, not to the weight of the entire argument, which involves the deontic premises and its plausibility assessment that, as argued above, does not lend itself to probability interpretation. Where applicable, this would then have to be brought forward as a part of the respective discussion.

Position (plan proposal) plausibility

It may now seem reasonable to suggest that the plausibility of the overall plan proposal PLPROP should be a function of all the argument weights AW of all the arguments brought forward in the discussion. So one might suggest a simple and straightforward aggregation function for this, such as

$$\text{PLPROP} = \sum_i^n (\text{AW}(i)) \quad \text{for all } n \text{ argument weights } \text{AW}(i) \quad \{8\}$$

This function should be taken with considerable caution, not only because there might be inconsistencies and other relationships among the various arguments. Given the specifications of the plausibility judgments and weights made, PLPROP according to this function will be a score on the original plausibility scale of -1 to $+1$. It would be appropriate to accept it as a reliable guide for decision only for a very simple interpretation of all the deontic claims of the “pro” arguments as additive “benefits” of the proposed plan, and all “con” arguments referring to “costs” to be subtracted from those benefits. In this, PLPROP can be seen as similar to an individual participant’s Benefit/Cost criterion, with all the caveats applying to the use of such measures as decision guides. Such caution is important not only because of the questions about simple additivity of positive and negative plan effects, but also because no discussion, even the most detailed and systematic one, can be relied upon to bring out all effects and concerns of plan proposals. Planning discussions must be considered to remain systematically incomplete, as discussed in more detail in Mann (2007).

More research and examination of the way individual argument weights should be aggregated into an overall individual position plausibility score is necessary for several reasons, such as those indicated. However, even such simple results as those produced by equation {8} can now be useful in the sense that if some decision-maker has made plausibility and weight assignments resulting in a strong negative overall proposal plausibility score, that decision-maker’s positive vote or decision on the proposal (or, of course, a positive overall result and a negative decision) would call for some additional explanation: were there additional factors not brought up in the discussion, that should be made explicit and discussed? In other words, it would now be easier—more transparent—to check on a decision-maker’s voting consistency with the assessments of the arguments that he or she had been able to substantiate in the discussion.

5. Group measures of assessment judgments and their use

It would be understandable and tempting at this point to use some overall group statistic of the individual plausibility results as a guide if not as an outright decision criterion—for example, the average of all proposal plausibility scores over a group. Such a measure would arguably already be better, because more deliberated, than using the result of an opinion poll that merely asks for support

or opposition to the proposal. However, this is neither advisable nor the intent of the proposed approach. The reasons are obvious: an average (mean) score would not properly reflect the distribution of such scores over the group: the same average can result from a majority of judgments clustered around the same score, as from two clusters at opposite ends of the scale. It would be meaningless if not outright dangerous to use such a measure without an accompanying measure of the range distribution of scores, or the variance, that would indicate the extent of agreement or disagreement about the proposed plan in the group.

The greatest potential value of the results of such a thorough evaluation lies in the *analysis* of the information, not in their use to guide or substitute for decisions. And the scrutiny and discussion of the distribution of results of individual arguments—premise and resulting argument plausibility, argument weight—will be more significant than a focus on overall proposal plausibility. Especially the analysis of large discrepancies in plausibility assessments of premises can help to identify real areas of agreement and disagreement. The latter can suggest issues where additional discussion and evidence might be needed, the former might show where sufficient agreement exists to make further repetitive discussion unnecessary—but it also might alert participants to the potential presence of uncritical “groupthink,” suggesting that some thought be spent on the “wrong question?” reminder. And plausibility assessment results hovering around the zero mark when the associated importance weights of the deontics are high should be considered a call for efforts to provide more and better information and research.

6. Procedural implications for planning discourse

The observations regarding the use of evaluation results suggest that adopting such an approach in the planning process will likely require some procedural adjustments. Obviously, the procedure will depend on the scope of a project, both as regards number of participants and/or affected parties, space, and time. For a quick demonstration sketch of the nature of needed adjustments, it will be sufficient to use an example of a project of small to medium scope, say a local planning issue to be decided by local government but requiring citizen input within a period of a few months.

General assumptions and agreements

It will be assumed that there is an entity such as a local government planning department that will organize the process. Its starting point could be either one or several more or less completely

worked-out plan proposal(s), or a problem statement, or simply a question that has the potential of evolving into a public controversy. The office would be in charge of preparing and presenting problem statements and plan proposals to the public, and for inviting and then compiling comments, questions, etc., from all sources—comments at meetings, editorials and letters to the editor, phone calls, emails, etc. These would be collected in chronological order in a “verbatim” file, preserving the original contribution for reference. Office staff would then analyze those contributions, “translate” them into the condensed format suggested above, (identifying issues, arguments, answers), sort and file these with the respective issues or questions.

In this condensed file, each argument will be entered only once (there will be no repetition of the same substantial point), but will be referenced to the various contributors that have made that same point, perhaps in different terms. This will result in a concise collection of pro and con arguments about the original proposal. Periodically, at least, but preferably on an ongoing basis, the office would publish, post or otherwise make available information on the current state of the discussion, ideally accompanied by suitable topic, issue, and argument “maps.”

Periodically, or when the flow of new contributions slows down, thereby suggesting that most main concerns have been voiced, an argument assessment “form” will be prepared. Participants will be invited to enter their assessments of premise plausibility scores and weights of deontics for the arguments, perhaps after having provided an initial offhand overall proposal plausibility score (for comparison with their “deliberated” results). The office will then calculate the argument weights and proposal plausibility for each participant individually, as well as any group statistics that have been agreed upon. The individual results will be returned to each participant; the group statistics will be posted on the “public” board together with the office’s analysis and suggestions for further discussion, need for more information, or modified plan proposals that respond to significant criticism or suggestions from the participants. This sequence can be repeated, if necessary until no more contributions, significant changes in information, or assessments occur. The results will be posted as a guideline for the “proper” governance body to make its decision according to established rules. Further considerations of implications for the planning process are discussed in Mann (2007).

*Technical tools for discourse support**Technical support*

The entire process of a planning discourse organized along the ideas described above can be carried out by traditional means with only minimal technical or IT support; it would just require some moderately trained personnel to do it. Obviously, providing technological support in the form of a website for collecting and displaying the material as it accumulates, for calculating the necessary results and for distributing these, would greatly streamline such a process. The very concept of such an interactive planning process with wide public participation is aimed at utilizing the potential of emerging information technology.

“Expert systems” for design and planning?

The task of organizing the support for such a process raises an interesting issue about the potential role of so-called “expert systems”—to what extent would “knowledge based systems” be able to assist and support such planning processes? Some brief consideration of the above description of the argument assessment approach (e.g., Mann 1995) should make it clear that traditional “expert systems,” whose superior abilities at gathering information (from actual observation and online search, etc.) and speed in computing consequences would seem to entitle them to make decisions quickly for their human users, will be entirely inappropriate here, for several reasons. The reliability of such systems rests on their ability to draw valid conclusions from incoming information and the in-built “knowledge base.” This requires that the knowledge base be internally consistent: free of contradictory statements.

But in planning and problem-solving, especially for unprecedented problems of the kind Rittel and Webber (1974) called “wicked problems” the “pro and con” discourse is, by definition, full of contradictory claims. The inference rules—even for “fuzzy” and informal, “plausible” inference-making—would not be able to draw meaningful, “valid” conclusions from such a data base so as to make independent decisions. Furthermore, the knowledge base itself, while it might of course contain the “textbook knowledge” for the respective domain, is inherently incomplete. The way a plan proposal will affect various parties, and how they react to those effects, cannot be known ahead of time. The complete pertinent knowledge base only *emerges* as the discussion proceeds, and may stray into entirely unanticipated knowledge domains.

These caveats regarding expert systems do not mean that such technology would be of no use at all for planning processes using the proposed argument assessment tools. For example, the system might store all contributions with all the premises brought forward

by human contributors and by the professional or scientific literature. Provided that the arguments and claims have been edited to the formal condensed version, it could then insert premises found into the argument structure patterns (inference rules) to construct arguments that have not been made thus far but should or might have been made given the available (or postulated) knowledge. It would then present these as contributions to the discussion, for evaluation by the human participants. It is here that the work of logicians—aiming at establishing what argument patterns are valid (or even just of some acceptable plausibility) and which patterns are useless—might be helpful in preventing the system from generating all the useless ones among all the variations of the basic pattern, and presenting them to human participants for evaluation. It should emphatically *not* be designed to make “its own” plausibility and especially weight assessments. It may be tempting to have it “simulate” a person or group’s assessment—but only upon having been given that person’s plausibility and weight scores for all extant information and then exploring, for example, what difference an additional argument might make in the person’s overall decision if the new information were considered plausible, important, or implausible, or itself inconclusive.

Overview: “mapping”

Another area where emerging technology can support this kind of planning process would be in the construction of “maps”—visual displays of the relationships among the topic, issues, and arguments. These are vitally important to provide human participants with a coherent overview of the discourse. There are several useful developments of such mapping technology software (not to be confused with GIS and similar technology): a few such examples are *Debategraph*³, *Dialogue mapping*⁴, *Compendium*⁵, *Araucaria*⁶, *Rationale*⁷ or *bCisive*⁸. So far none of these, to my knowledge, has been applied to the special aspect of planning arguments nor adapted to the level of detail of argument analysis that would facilitate argument evaluation as described above. Some formats for topic, issue, and argument maps that would better match the described argument assessment approach are suggested in Mann (2007) but have not been developed as program applications yet.

More work on these problems is needed. Arguably, most of the discussions and controversies humanity argues about have to do

³ Debategraph: <http://www.debategraph.org/>

⁴ Dialogue mapping: <http://cognexus.org/>

⁵ Compendium: <http://compendium.open.ac.uk/institute>

⁶ Araucaria: <http://araucaria.computing.dundee.ac.uk/>

⁷ Rationale: <http://rationale.austhink.com>

⁸ Bcive: <http://bcive.austhink.com>

with issues of what we ought to do—in planning, problem-solving, policy-making, politics. Theory, research, programs and technology development should devote more effort to this problem of evaluation of the merit of planning arguments.

7. Conclusion

Arguably, most of the discussions and controversies humanity argues about have to do with issues of what we ought to do—in planning, problem-solving, policy-making, and politics. While the technology of communication, information processing and computing has made great strides during the past few decades, the tools for the evaluation of all this information, and especially of the arguments being put forward in such planning and policy-making discussions, have not kept pace with this development. The ideas, problems, and proposals sketched out in the preceding should be seen as some first steps, suggestions and an invitation for discussion of a comprehensive program of research⁹, development of technology as well as procedural tools, and experimental application on the topic of more systematic and transparent evaluation of design and planning arguments.

References

- Ackoff, R, Churchman, C.W. & Smith Jr., N. (1954). “An Approximate Measure of Value”. *Journal of the Operations Research Society of America*, 2(2).
- Galle, Per. (1996). “Design rationalization and the logic of design: a case study.” in: *Design Studies* 17(3).
- Galle, Per. (1997). “Towards a formal logic of design rationalization.” in *Design Studies* 18(2).
- Gordon, Thomas & Karacapilidis, M. (1999). “The Zeno Argumentation Framework.” *German Artificial Intelligence* 3, 1999.
- Gordon, Thomas *et al.* (2001). “Zeno: Groupware for Discourses on the Internet,” *Künstliche Intelligenz*, Vol. 2/01.
- Habermas, Jürgen. (1968). *Erkenntnis und Interesse* Suhrkamp, Frankfurt. (English: *Knowledge and Human Interests*, Boston, MA: Beacon Press, 1971).
- Habermas, Jürgen (1973). *Legimitationsprobleme im Spätkapitalismus*, Suhrkamp, Frankfurt .

⁹ A more detailed discussion of a tentative research program is presented in Mann (2007).

- Habermas, J. (1984). *Theorie des kommunikativen Handelns*. (English, translated by Thomas McCarthy, Boston, MA: Beacon Press.)
- Hall, James. (2005). *Tools of Thinking*. Chantilly, VA: The Teaching Company.
- Kunz, W. & Rittel, H. (1970). "Issues as Elements of Information Systems." Working Paper No. 131, Institute of Urban and Regional Development, University of California, Berkeley.
- Musso, A. & Rittel, H. (1969). "Über das Messen der Güte von Gebäuden" in *Arbeits-berichte zur Entwurfsmethodik*, Stuttgart: Krämer Verlag.
- Mann, Thorbjørn. (1977). *Argument Assessment for Design Decisions*, Dissertation, University of California, Berkeley.
- Mann, Thorbjørn. (1995). "Expert Systems for Design and Planning: Requirements and Expectations," *Proceedings*, International Conference on Engineering Design (ICED) 95, Prague, Czech Republic.
- Mann, Thorbjørn. (2007). *The Fog Island Argument*, XLibris, (In German: *Das Planungsargument*, E-book, Ciando, 2008)
- Popper, Karl R. (1968). *The Logic of Scientific Discovery*. New York: Harper & Row.
- Rittel, H. & Webber, M. (1974). "Dilemmas in a General Theory of Planning," *Policy Sciences* 4.
- Rittel, H. et al. (1972). *Intensivere Nutzung der räumlichen Kapazität im Hochschul-bereich*, Project for the German Bundesministeriums für Bildung und Wissenschaft, Bonn, Project report. Heidelberg: Studiengruppe für Systemforschung.
- Rittel, H. (1972). "On the Planning Crisis: Systems Analysis of the 'First and Second Generations'." *BedriftsØkonomen*. #8.
- Rittel, H. (1977). "Structure and Usefulness of Planning Information Systems", Working Paper S-77-8, Institut für Grundlagen der Planung, Universität Stuttgart.
- Rittel, H. (1980). "APIS: A Concept for an Argumentative Planning Information System". Working Paper No. 324. Berkeley: Institute of Urban and Regional Development, University of California.
- Rittel, H. (1989). "Issue-Based Information Systems for Design". Working Paper No. 492. Berkeley: Institute of Urban and Regional Development, University of California.
- Toulmin, S. (1959). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Zarefsky, David. (2007). *Argumentation: the Study of Effective Reasoning*. The Teaching Company.
- Zenker, Frank. (2009). On the Reconstruction and Evaluation of Conductive Argumentation. Dept. of Rhetoric, Communication & Media Studies. University of Copenhagen, Denmark (invited talk.)

Zenker, Frank. (2010). "Deduction, Induction, Conduction. An Attempt at Unifying Natural Language Argument Structures." *Proceedings*, 2010 conference on argumentation, Wake Forest University, Wake Forest, N.C., USA.