

## ARGUMENTATION, STORIES, AND GENERALIZATIONS: A COMMENT

**William Twining\***

### **Abstract**

The underlying theory of the software program described in “Sense-making Software for Criminal Investigation” (above 000) complements modified Wigmorean analysis (MWA). Both adopt a qualitative rather than a quantitative approach. MWA is broadly compatible with the kind of logic involved, including abductive inference to the best explanation and the idea of defeasible argumentation. Both approaches are mainly valuable as aids to thinking, especially constructing and evaluating arguments, rather than as methods of presenting them in order to persuade. Both approaches can be applied at different stages of criminal investigation (and more broadly of legal processes), but the specific device of Wigmore charts (one part of MWA) is more useful in hypothesis-testing and discarding than in hypothesis formation, which typically requires imaginative reasoning. Crombag et al.’s *Anchored Narratives* and MWA have similar theoretical assumptions, except that MWA gives a radically different account about the relationship between stories, generalizations and argument. The proposed program has considerable promise, but before it can be of positive practical value in police investigation, more attention needs to be given not only to the obvious dangers of using stories and generalizations in this context, but also about what positive guidance can be given to mitigate these dangers. There is, however, an unresolved tension between the simplifying tendencies of formalized computer programming and the tendency of MWA to emphasize the complexities of practical inferential reasoning and argumentation in legal contexts.

### *Introduction*

In their paper “Sense-making Software for Crime Investigation”(hereafter ‘Sense-making’), Henry Prakken and his colleagues describe a project that is both intellectually and practically ambitious: it

---

\* Quain Professor of Jurisprudence Emeritus, University College London and Visiting Professor, University of Miami School of Law. I am grateful to Terry Anderson, David Schum, Peter Tillers and participants at the Cardozo Conference in January 2007 for helpful comments and suggestions.

seeks to synthesize ideas about inferential reasoning and narrative in a single formal model of reasoning about evidence; its goal is to produce software that has potentially wide practical application in criminal investigation (and possibly related types of enquiry) especially at the stage of hypothesis formation.

The aim is to produce software “that is both natural and rationally well-formed” (“Sense-making” at 000 < p.1 abstract >), that is to say that it will be closely related to existing best practice in police investigation, it will be compatible with other systems, and it will draw on advances in informal logic, artificial intelligence, and modern evidence theory. The basic framework is argumentation. The intention is admirable and the general project seems very promising. It is especially welcome that the focus is on making the investigators’ reasonings explicit and that the project aims to combine stories and arguments within a framework of argumentation using concepts that are actually used in evidential reasoning.

The purpose of this comment is to make some constructive suggestions for further refining this potentially useful tool and its underlying theory in light of an expanded conception of modified Wigmorean analysis (hereafter MWA). It is written from the standpoint of a theorist of evidence, who is sympathetic both to the objectives of the project and to most, but not all, of its theoretical assumptions, but who is concerned that in their efforts to design support tools that are accessible, usable, and “natural” the authors may gloss over some of the complexities and dangers associated with stories and generalizations in argumentation based on evidence, especially in police investigation.

In the past I have been cautious about this kind of software for three main reasons:

- (i) there is a danger that the software will oversimplify the contexts and the tasks involved (in this case in criminal investigation) and in the process build in dubious short-cuts and simplifications;
- (ii) there is a danger that a program designed to be an *aid* to thinking will be used in practice as a *substitute* for thinking (Leary 2003, 2004)
- (iii) If such software adopts an approach that uses generalizations and stories, the users need to be warned about the dangers. In reasoning about questions of fact, both stories and generalizations are necessary, but dangerous.

In considering how these pitfalls can be avoided, I shall say little about formalization or methods of representation. Rather I shall concentrate on the treatment in “Sense-making” of the concepts of stories and generalizations, their roles and relations in enquiries about questions of fact in legal contexts, and the special dangers attending their use in police investigation.

*The New Evidence Scholarship and Anchored Narratives.*

“Sense-making” contrasts two theories that it seeks to synthesize: New Evidence Theory (exemplified by Wigmore and his successors) (hereafter NET) and the theory of Anchored Narratives (hereafter ANT) developed by Crombag, Van Koppen and Wagenaar (Crombag et al, 1992, 1993):

“Both approaches stress the importance of empirical generalizations in evidential reasoning but they differ on their precise role. While NET regards generalizations as the “glue” in evidential arguments from evidence to hypotheses, ANT regards them as “the anchors” of evidential stories in the available evidence.” (Sense-making at 000 <p.3>)

I am uneasy about this juxtaposition for two main reasons; first, the brief account in the paper, though generally correct, does not mention some important features of modified Wigmorean analysis (hereafter MWA), including its account of the relationship between arguments, generalizations, and stories.<sup>1</sup> Second, I suggest that the theoretical framework of *Anchored Narratives* involves two serious flaws and, if these are corrected, there is not much difference between the two approaches that the authors of “Sense-making” seek to synthesize.

There are several common misunderstandings about MWA. First, there is a tendency to talk of “Wigmore charts” and “the Wigmorean method” without clearly differentiating between Wigmore’s classic statements (Wigmore, 1913, 1913a, 1931, 1937) and subsequent refinements and developments (footnote <1> above) It is as anachronistic to compare Wigmore’s first efforts with the modern approach as it is to equate Bayes’ original account of his thesis with modern Bayesianism.

In respect of law, MWA differs from Wigmore’s original account in the following important respects: First, instead of treating the contested jury trial as the paradigm arena of evidential reasoning in law, MWA adopts a total process model of civil and criminal litigation starting with some initial event or claim or accusation and continuing on to include all important pre-trial and post-trial decisions, including, for example, negotiations, settlements, sentencing, decisions to appeal, appeals, parole, and predicting

---

<sup>1</sup> In this comment MWA is treated as one part of the broader movement labeled “New Evidence Theory” (NET), (see Twining (1991), Jackson (1996), Park (2001)). MWA is expounded in the second edition of Anderson, Schum and Twining, *Analysis of Evidence* (2005)(hereafter *Analysis*). For more detailed discussions see also Schum (1994) (1999), (2003), Twining (1985), (2002), (2006), Tillers and Schum (1991), (1998). “Sense-making” relies on the first edition of *Analysis* (Anderson and Twining, 1990, 1998), but most of the points made here about stories and generalizations have been in print for some years. On generalizations see especially Anderson (1999); on stories and generalizations and how they are connected, see the papers collected in Twining (2002) Ch. 12-16, (2006) Ch. 9-13.

future dangerousness in death row cases. MWA emphasizes that questions of fact, evidentiary issues, and inferential reasoning can be relevant to nearly all important decisions in legal process, not just in trials. We agree with “Sense-making” that modified MWA is relevant to all stages of investigations.<sup>2</sup>

Second, and closely connected, MWA emphasizes clarification of standpoint as an essential preliminary for analysis. In teaching, we emphasize that a Wigmore chart is typically a picture of the chart-maker’s own argument relevant to a specified ultimate probandum or question of fact at a given stage in a particular process on the basis of a specified data base or body of source material. As a typical case proceeds the roles, the vantage points, and the available and usable data can change at each stage. It is important to differentiate between stages within an investigation and in this respect distinctions between hypothesis formation, hypothesis testing and hypothesis elimination are especially significant. The general approach is relevant to all three, but actual charts are mainly useful in respect of arguments supporting or negating a given hypothesis or hypotheses.<sup>3</sup>

Third, Wigmore treated his method and “the narrative method” as alternatives. MWA treats them as complementary. MWA acknowledges that there are puzzles about the relationship between stories and argument and we have sought to give an account of this, which will be explored below. However, we resist any suggestion that stories can be included within Wigmore charts.<sup>4</sup>

Fourth, MWA compares and contrasts a number of devices for “managing” facts<sup>5</sup> and marshalling arguments so that a case can be seen as a meaningful whole – chronological tables, trial books, Wigmore charts, and stories all perform this function in different ways. (*Analysis* (2005) Ch. 6 and 12A). Each device has its uses and limitations at different stages of legal processes. Stories and charts have other roles in addition to this.

---

<sup>2</sup> In the 1937 edition of *The Science of Judicial Proof*, Wigmore included a short appendix on investigation, but it was not developed very far. Contrast *Analysis* Ch. 2.

<sup>3</sup> “Sense-making” is mainly concerned with hypothesis formation, and the analysis of the *King* assumes a fairly advanced stage in this process (<p. 15>). However, the proposed software should be sufficiently flexible to be usable at some other stages.

<sup>4</sup> In “Sense-making” it is stated that NET “does not incorporate this in the use of Wigmore charts.” (00 =p.3) If this means that stories cannot be included within Wigmore charts, this is correct, because the only relations depicted in such charts are logical, not temporal or causal. (see below) If this means that MWA says nothing about the role of stories at different stages in legal process, that is incorrect (see references in n.000 <1> above).

<sup>5</sup> In the United Kingdom the general approach is sometimes referred to as “fact management”.

Fifth, MWA makes a clear distinction between macroscopic and microscopic analysis. The method has two aspects. Macroscopic analysis provides a simple way of marshalling, and structuring a complex argument so that it can become manageable. Microscopic analysis is used to subject *selected* key phases of the argument to detailed scrutiny within the framework of the case as a whole (see below)

Sixth, MWA distinguishes between several types of generalization and between different roles of generalizations at different stages of legal processes.<sup>6</sup> In the present context it is especially useful to distinguish between scientific, experience-based,<sup>7</sup> common-sense, and case-specific generalizations. (*Analysis* 265-76)

Seventh, MWA deals with investigation, which involves hypothesis formation, hypothesis testing, and hypothesis elimination. (*Analysis* Ch.2) In criminal investigation the primary manifest function is to collect potential evidence for and against one or more hypotheses. Asking good questions, identifying potentially fruitful lines of inquiry, constructing rival hypotheses, selecting potentially relevant items (“dots” or trifles”) from a mass of information, “joining the dots”, and identifying what further information is needed are among the skills required. Stories, situation-types, familiar fact-patterns, and generalizations can all play significant roles in these operations.<sup>8</sup>

Apart from conflating MWA with Wigmore’s original accounts, there are several common misapprehensions about Wigmorean method. First, it is frequently said that the chart method is too complex to be of practical use.<sup>9</sup> This is to confuse the solution with the problem. MWA is a *simplifying* method for dealing with *complex arguments* based on *mixed masses of* evidence. It is an aid to dealing with *complex cases*. (Malsch and Nijboer 1999)<sup>10</sup> The logic is simple and binary (all relations take the

---

<sup>6</sup> In *Analysis* (2005) the idea of standards of proof is extended to the more general category of standards for decision, including, for example, decisions to arrest, search, charge, prosecute, plead guilty or not guilty, all of which are closely connected to different standpoints at different stages of the total process. Several of these decisions can be involved in a police investigation.

<sup>7</sup> E.g. experience-based “knowledge” of the police about local patterns of vehicle theft (see below 000).

<sup>8</sup> MWA also modifies or diverges from Wigmore’s original account in a number of ways which are not directly relevant to this comment: for example, the symbols are simplified, problems of evaluation (including probability theory) are discussed at length, as is the relationship between principles of proof and the law of evidence.

<sup>9</sup> E.g. Murphy (2000)1.n., Roberts (2002) discussed in *Analysis* (2005) at 123-4, 140-2..

<sup>10</sup> To date nearly all examples of attempts at formalization deal with simple cases, which in Wigmorean terms require a key-list of less than 50 propositions. In the exercises in MWA that we set our students I prescribe an artificial limit of 500-600 propositions, which usually means that they have to do selective microscopic analysis of

form “tend to support” “tend to negate”). Macroscopic analysis provides a simple way of marshalling, and structuring a complex argument so that it can become manageable. Unless the law defining the material facts (the ultimate and penultimate probanda) is unclear, constructing the top of the chart is typically simple, quick and time-saving. It accords with best practice in both simple and complex cases.

Once the top of the chart is settled, the next step is to pause and construct provisional theories of the case for each side. This requires both logic and imagination. Stories, scenarios, and themes can all be useful in the process. This stage of strategic thinking also serves to simplify the task in two ways: it encourages the analyst to think in terms of elegant, streamlined, economical arguments; it also enables the analyst to identify and select propositions that play a critical role in the argument as a whole – what some litigators call the “jugular(s)” – so that these become focal points of attention. Often a case can be won or lost by establishing or casting doubt on a single proposition. Nearly all such propositions are to be found at or near the top of the chart<sup>11</sup> – this enables the analyst to be highly selective about which phase(s) of the argument and which evidence should be the subject of intensive microscopic analysis, which can, of course, be intricate, complex and time-consuming. But again the difficulty and the complexity lie in the subject-matter not in the method. The logic is the same and the number of patterns of relationship is limited. MWA saves the analyst from floundering around in a mass of disorganized and confusing bits and pieces.

Another misapprehension about MWA is that its main value is as a form of *representation* of arguments. In teaching law students, I emphasize a distinction between preparation and presentation of arguments. The main value of a Wigmorean approach is to assist in the *construction* of an argument – especially to clarify with precision exactly what the argument is, to identify weak points on either side, and to consider whether they can be strengthened or destroyed. The main skill is to construct the strongest potential argument for each side in a single dialectical process. Once the argument has been constructed, concerns about communication and persuasion raise different issues. The Wigmorean analyst need have no illusions that the most cogent argument is ipso facto the most persuasive, but she probably does need to have faith that cogency plays some role in persuasion. The proposed software program under consideration is intended as an aid to *rational* police investigation. (<pp1-2 >) Like MWA its chief value is heuristic; communication within a team or organization is important, but secondary. Wigmore charts

---

parts of the argument. A partial exception is Kadane and Schum (1996) (a book length analysis of the *Sacco-Vanzetti* case).

<sup>11</sup> An important exception is supporting or undermining the credibility of a key witness (ancillary evidence).

have limited value as a means of communication. This is partly because they involve the use of symbols, which like equations and formulas are off-putting to the uninitiated. It is also because to date most users of the method keep their key-list and charts separate – important, perhaps necessary at the stage of argument construction.<sup>12</sup>

A third misapprehension is that MWA can incorporate stories or causal explanations or intellectual protocols *within* the charts.<sup>13</sup> Part of the confusion arises from ambiguities in the use of arrows. A Wigmore chart is a picture of the relations between all of the propositions that form a part of an argument. The relations are only of two kinds – tend to support ( $\uparrow$ ) and tend to negate ( $\rightarrow$ ). But arrows are also used in other contexts to indicate temporal flows (“and then” ... “And then”) or causal relations (propter hoc/resulting in) or next steps (as in an algorithm) or relationships or connections (e.g a kinship system) or merely a vague “some connection”. In our experience the most common error that students make in their first efforts at charting is to mix in temporal or causal relations with logical ones. Unless trained to differentiate between these types of relations and to use different symbols to depict them, police investigators could easily fall into this trap.

### *Anchored Narrative Theory*

Nearly twelve years ago, in a quite gentle review, I praised the authors of *Anchored Narratives* for transporting Anglo-American evidence theory to the Netherlands and applying it to detailed analyses of the facts of some Dutch cases. One of the achievements of this important book was to show in detail how some aspects of common law evidence theory “travel well” across different procedural systems and legal cultures. It also contained some significant psychological insights. But I criticized their account of the underlying theory. (Twining, 1995 reprinted in (2002))

First, I suggested that they conflated empirical generalizations and prescriptive rules and made a quite unwarranted jump from some empirical psychological findings to over-simple and very dubious

---

<sup>12</sup> Palmer (2003) illustrates one way of inserting propositions within nodes on a chart. This is quite useful for purposes of communication with relatively simple charts, provided it does not clutter the presentation. However, one important skill involved in argument construction is drafting each proposition with precision and, in the case, of significant generalizations, constructing the least vulnerable formulation. Constructing a key-list is logically (and, for some temporally) anterior to constructing a chart, which is a chart of the relations between all the propositions on the key-list.

<sup>13</sup> See above n. <4>.

“universal rules of evidence”.<sup>14</sup> There is a fundamentally important distinction between empirical generalizations and prescriptions, and one cannot assume that rules in the prescriptive sense can be built so simply on generalizations. Of course, in practice it is often hard to keep fact and value separate, and this is one of the dangers of relying on allegedly empirical generalizations (see below).

More important is the claim that in argumentation about questions of fact in legal contexts the main anchors of stories are generalizations. (Crombag et al, 1993 37-42, especially Figure 2.1) This is an error. The main anchors for a believable (“plausible” is ambiguous)<sup>15</sup> story in this context must always be particular evidence. It is a necessary condition of a believable story that each significant element is anchored in case specific evidence about a particular past event. Just as naked statistical evidence is easily trumped by particular evidence in the present case, so using generalizations as anchors for the truth of a story, is at best a last resort. Bennett and Feldman give the game away when they claim that an important function of stories is to fill in gaps in the evidence. (Bennett and Feldman 1981 pp.33, 65; discussed Twining (2006) 307-8) An advocate who uses generalizations in this way has a weak case.

“Sense-making” appears to avoid both these errors. The authors make it clear that they are talking about empirical generalizations rather than rules; and they quite correctly anchor their analysis of the *King* case in the testimony of three witnesses (<pp. 6-8>). If one remedies these two errors in ANT, it is not clear to me how the underlying theory differs from MWA, except that the authors of “Sense-making” seem to include generalizations in an extended meaning of “anchors”. They distinguish between external and internal anchoring. The former relates to the quality of the story (e.g. coherence, consistency, completeness), and they usefully suggest that a story is stronger if connections between the events in a story are causal as well as temporal (000 <p3>). They state that external anchoring depends on particular

---

<sup>14</sup> To take one specific argument from the common law: there is a great deal of empirical evidence showing that in given circumstances, eyewitness identification evidence is generally unreliable. To the best of my knowledge, no common law jurisdiction has introduced a rule requiring corroboration of this kind of evidence. The reason often given for this is that there are numerous exceptions to the generalization. But we have enough empirical evidence now to produce a more nuanced rule that could specify under which circumstances this kind of evidence is especially unreliable and requires corroboration. Why is there no rule? One reason is that we are dealing with a huge variety of combinations of situation and so any rule is likely to be over- or under-inclusive or both. Another reason is that the circumstances often differ in respect of continuous variation and are not susceptible to bright line rules. A third is a reluctance to proliferate exclusionary rules. Other considerations apply to differences between procedural systems \_\_\_ jury/non-jury; criminal/ civil; and so on. (Twining (2006) at 000).

<sup>15</sup> [“Plausible” may mean believable or it may mean persuasive or attractive (see below 000) ].

evidence, but that commonsense generalizations have a role. MWA suggests that in this context the main role is to warrant inferential steps from particular to particular and only in exceptional cases (e.g. judicial notice, an absence of particular evidence, a generally incredible story)<sup>16</sup> to lend direct support for a relevant fact. This is different from ANT, which suggests that generalizations are the main anchors for stories.<sup>17</sup>

Insofar as they substitute particular evidence for generalizations, the authors of “Sense-making” are closer to MWA than to ANT. So far as “internal anchoring” of stories is concerned, the accounts are very similar, although the language is different. In the 2005 edition of *Analysis* there is a “protocol for assessing the plausibility, coherence, and evidentiary support for a story” (*Analysis* at 281-2). That protocol includes all of the elements covered by the concepts of internal and external anchoring and a number of others. So far as external anchoring is concerned, the most important kind is particular evidence. This story-testing protocol might be refined and adapted for use in the kind of program under consideration.

*Concepts: story, generalization, and story-type distinguished*

At first sight the concepts of “story” and “generalization” are easily distinguished. Following Paul Ricoeur, one can define a story as ““a narrative of particular events arranged in time and forming a meaningful totality.”(Twining (2006) 223 adapted from Ricoeur (1981)).<sup>18</sup> Defined precisely, a story has three elements: particularity, temporality, and coherence or unity. Each of these elements is lacking in commonsense and scientific generalizations of the kind that form “the glue” in inferential reasoning. Such a generalization can be defined as “a general proposition claimed to be true which is used implicitly or explicitly to argue that a conclusion has been established”. (*Analysis* 383) Generalizations are general, stories are particular; generalizations are single (typically simple) propositions, stories are expressed in many statements or propositions; the idea of coherence or unity of stories does not apply to single, simple

---

<sup>16</sup> If a story is impossible in some material way, judged by one’s stock of beliefs (she flew off on a broomstick, she was transformed into a fairy godmother), then a generalization can be used to reject a story; but if a story supported by credible particular evidence is unlikely, even bizarre, particular evidence usually trumps the story, for unlikely events do happen.

<sup>17</sup> “[A]lthough the proponents in ANT focus on the story-based perspective in their choice of both wording and research background, several of their central claims have a more argumentative than story-based flavour” (Sense-making at p. <3>). Insofar as this correct, the differences between ANT and MWA are quite small.

<sup>18</sup> On this view causality is an important, but not a necessary, feature of stories. What kinds of connectedness or coherence gives unity to stories other than causal relations is a matter of debate that I do not pursue here.

generalizations; in this context, generalizations are claimed to be empirically true, but no such claim is made for many stories;<sup>19</sup> stories give accounts of change over time; temporality is not a necessary feature of generalizations, but, of course, some generalizations are about patterns of change.

These seemingly discrete concepts are connected in quite subtle ways. Indeed, there is an intimate relationship between them. The “plot” of a story can be summarized in a single sentence: “Poor girl loses slipper and wins prince”. Such a proposition is easily generalized: “Poor girls who lose a slipper, (often/usually/ always) win a prince”. Closely related to this, the point or moral or *ratio decidendi* of a story is often expressed as a general normative proposition.<sup>20</sup> The point or moral of the Cinderella story could be rendered in many different ways: “If you want to catch a prince, lose a slipper” (prudential); “beautiful poor girls are more deserving of the love of a prince than rich ugly sisters” (normative, egalitarian); “determined princes find the right girl” (empirical?); “love conquers all” and so on. Part of the power of stories is the indeterminacy of their point or moral. Conversely, a particular story may be used to illustrate, explain, concretize, or even give evidentiary support to a generalization.

There is a useful concept that lies at an intermediate point between stories and commonsense generalizations: a generalized story or story-type, sometimes referred to by such terms as “plot” or “scenario”.<sup>21</sup> In ordinary usage both of the latter terms are used loosely to refer either to particular stories (the plot or scenario of this film) or to story-types (the classic detective story, the love triangle, even the Cinderella story). We shall see below that one of the dangers is that it is easy to treat a story in a case as plausible or even true because it is familiar or reassuring or pleasing in some other way, not because it is empirically true. (See further below).

Generality and particularity are relative matters. Some stories fit a standard plot or pattern; some generalizations are case-specific (John is a smoker; Cinderella is careless with her shoes). A story may be told in quite abstract categories, but still be an account of a particular event; generalizations can be quite detailed and, in that sense, quite particular. Common lawyers are familiar with the intimate relationship

---

<sup>19</sup> Of course, in litigation the truth of stories put forward as accounts of “the facts” is in issue.

<sup>20</sup> This theme is explored at length in “The *Ratio Decidendi* of the Parable of the Prodigal Son” in Twining (2002) Ch. 16, reprinted in Twining (2006) Ch.13.

<sup>21</sup> According to the OED the primary meaning of “scenario” is “a sketch or outline of a plot or play, ballet, novel, opera, story etc.”. This usage is particular, but the dictionary recognizes weakened or loose usages, which do not distinguish clearly between general and particular, e.g. “a scientific model or description intended to account for observable facts” (ibid.). {Tillers “a series of events causally linked in time” [ck T ad S](#)}

between the specific facts of a particular case and the general rule (*ratio decidendi*, holding) for which the case is cited as authority. How the facts are categorized determines the generality of the rule for which the case is being made to stand.<sup>22</sup>

Fourth, both generalizations and stories form part of our so-called, allegedly shared “stock of knowledge” or “stock of beliefs” — that undifferentiated bouillabaisse of beliefs, scientific truths, prejudices, and impressions, which is loosely referred to as common sense or general experience.<sup>23</sup> In short, in this context generalizations and stories are frequently derived from the same source.

To sum up: stories and generalizations are separate concepts, but in the context of argumentation and enquiry about disputed or unsettled questions of fact in legal processes, they are intimately connected from both logical and psychological points of view.

#### *Multiple roles for stories, story-types and generalizations*

“Sense-making” shows convincingly that MWA is basically compatible with abductive inference to the best explanation and defeasible argumentation from particular to particular in a structured argument. (See generally Prakken (1997) Both approaches give quite similar accounts of the role of stories and generalizations in argumentation, but they are expressed differently. At a general level both emphasize the role of generalizations in warranting each inferential step from particular to particular (“the glue”) and the role of stories in “making sense” of an alleged past event. However, MWA goes further in suggesting that both stories and generalizations have other roles in argumentation about particular past

---

<sup>22</sup> This can be rendered by the proposition  $X=X=X$ : X happened (statement of the facts); if X happens, then what? (question of law); Whenever X happens then Y (legal rule). X is a constant. The less detail in X the broader the rule. E.g. Twining (2006) 405-6. The classic account by Julius Stone of “a ladder of abstraction” indicates the close connection between how the facts of a case are characterized and its *ratio decidendi* (Stone, 1959). Karl Llewellyn’s notion of “type-fact situations” is very similar. (Llewellyn, 1960)

<sup>23</sup> If I may quote myself: “A ‘stock of knowledge’ does not consist of individual, empirically tested, and readily articulated propositions; rather both individually and collectively, we have ill-defined agglomerations of beliefs that typically consist of a complex soup of more or less well-grounded information, sophisticated models, anecdotal memories, impressions, stories, myths, proverbs, wishes, stereotypes, speculations, and prejudices. Fact and value are not sharply differentiated. Nor are fact, fantasy, and fiction. Nor can one take for granted either consistency or coherence within an individual’s or a society’s ‘stock of knowledge’”. (Twining (2002) at 454-5.)

events and in emphasizing that some of these roles are illegitimate, at least in the context of legal processes.<sup>24</sup>

MWA distinguishes between legitimate and illegitimate uses of generalizations and stories in constructing, evaluating, and communicating arguments based on evidence about disputed or puzzling questions of fact. Without attempting a comprehensive statement of a view which needs further development, one can at least suggest some legitimate roles for both stories and generalizations.

The main legitimate roles for stories reflect the three elements in Ricouer's conception of a "story". First, the particularity of stories usually makes them more vivid, more memorable, and easier to grasp than more abstract forms of discourse. So they are especially useful in *presenting and communicating* a message (including an argument) to an audience. Second, stories have a unity. An account of an alleged event constructed as a narrative is one way of grasping "the case as a whole". Stories are not unique in that respect, for chronological tables and Wigmore charts claim to have similar functions. Third, the temporality of stories, as with chronological tables, can at least give an account of how different elements are related in time and, as "Sense-making" suggests, temporal links are often also causal links (but one needs to beware of the *post hoc, propter hoc* fallacy). So a story often gives an account not only of what happened, but also explains why it happened. Even in cases where motive does not have to be proved as a material fact (penultimate probandum), it is widely recognized by litigators, backed by psychological research (e.g. Hastie (ed.) 1983, Hastie, Penrod and Pennington, 1993), that jurors want to learn about motives.

Thus when it is claimed that stories help us to "make sense" of a case (e.g. White, 1985) this is because they are seen as readily understood and "natural" means of giving unity to our understanding of a case and explaining the events in broadly causal terms.

We are agreed that generalizations are needed as the "glue" for inferential reasoning. However, they also can play other roles at different stages of rational inquiry in legal processes. Because MWA differentiates between different stages and standpoints in legal processes, including different stages in a typical investigation, the question arises: what is the role of stories and generalizations at each stage? This is a question that requires further exploration, but in the present context let me suggest that a general

---

<sup>24</sup> The idea of "legitimacy" in this context is problematic. Suffice to say here that any moves or ploys that violate canons of ethics, substantive law, or (in the context of a trial) exclusionary rules is "illegitimate". There is plenty of scope for differences of interpretation of such rules and for what, more generally, constitutes "cheating" or unfairness in argumentation, but these cannot be explored here. (See Twining 2006 p.321 n10). |

answer is that they can have multiple roles and that some of these roles are shared because of the close relationship between stories, story-types, and different kinds of generalization. For reasons of space, a single example will have to suffice here.

It is clear that all three can play a role in the process of hypothesis formation (loosely referred to as abduction or imaginative reasoning). For example, in Conan Doyle's story "Silver Blaze" (Doyle, 1893), Sherlock Holmes famously seized on one "trifle" (simple fact) to explain the disappearance of a racehorse and to identify the prime suspect. Given the bare facts that during a particular night the horse had disappeared from stables that were guarded by a fierce dog, Holmes seized on the fact that the dog did not bark in the night and inferred that the thief/culprit was well-known to the dog ("dogs only bark at strangers" \_\_ generalization). This eliminated the two main suspects, a suspicious stranger in the vicinity and the trainer from (or someone else connected with) the nearby rival stables. We might reconstruct the next stages in Holmes' reasoning as follows: six non-strangers had opportunity, but four can be quickly eliminated as potential suspects for other reasons; that leaves the groom and the owner. Although not fully articulated in Conan Doyle's account, in considering possible motives, Holmes could have drawn on a stock of racing story-types: the culprit had backed a rival horse (trainer or owner); the horse was heavily insured (owner); the culprit was bribed by a third party, e.g. a rival owner or a bookie (trainer); revenge (trainer) and so on. These hypotheses could equally well take the form of particular stories, story-types, or generalizations. Eventually, Holmes settles on the trainer, on the basis of additional information. This is one possible reconstruction (there are other possibilities) of Holmes's imaginative reasoning to construct a hypothesis about the disappearance of the horse.<sup>25</sup> It illustrates how generalizations, particular stories, and story-types can all help in the construction of a hypothesis about identity and in due course, after further inquiry guided by these initial steps, of a coherent and believable story about the event as a whole.

To sum up: we agree that argument, generalizations and stories can be combined and that they complement each other. However, they have more varied roles than "Sense-making" suggests and they interact with each other in different ways at different stages of legal process.

---

<sup>25</sup> The facts in the original story are more complex, as is Holmes' reasoning. In the story attention had been diverted from the key clues as to the identity of the thief (the dog did not bark, the stable boy had been drugged) by the presence in the vicinity of a suspicious stranger and the fact that, on the day after the horse disappeared the trainer (Straker) had been found dead, killed by a vicious blow to the head. In Conan Doyle's rendering, Straker removed the horse in order to incapacitate it, the horse killed Straker in self-defense, and the rival trainer found and (illegally?) detained Silver Blaze. Thus the horse was the killer and the rival trainer the thief!

## *Dangers*

As was stated at the start, I am cautious about software programs of this kind because the disciplines of AI and argumentation theory support a tendency to over-simplify and because of the temptation to treat programs designed as aids to thinking as if they are substitutes for thinking. The authors of “Sense-making” emphasize that the user of the system “is responsible for carefully testing the quality of his stories and arguments” (< 7>) and that they plan to “inform the user about the dangers involved in relying on stories and generalizations.” (<8>) However, one wonders whether this goes far enough. In particular, there are difficult and as yet unresolved problems about what kinds of generalization and stories police investigators should be encouraged and helped to use. A central dilemma for this kind of software design is how, on the one hand, to make the program comprehensible, “user friendly”, and “natural” (in the sense of being close to existing ways of thought) without encouraging or leaving the door open to the kinds of corner cutting, prejudice, bias, speculation, and “cheating” to which both stories and generalizations are prone.<sup>26</sup>

A central theme of my own writing in this area has been that both stories and generalizations are “necessary but dangerous”. (e.g. Twining (1999), (2002) Ch.12-16, (2006) Ch.9-13, *Analysis* Ch.10). It would be inappropriate here to go over this ground again. I hope that these ideas will be helpful to the authors in constructing their warnings to users. However, some of their remarks in “Sense-making” and the authors’ own use of generalizations suggest that there is still some theoretical groundwork to be done and that the difficulties involved may have been underestimated.

My position is that stories and generalizations are both *inherently problematic* in this kind of context. In respect of stories, it is generally recognized that stories have a fundamental human appeal, for a variety of reasons in addition to their particularity, vividness, and the leeways for interpretation of point or significance. To put it succinctly: in the present context we are concerned with stories as means to making good judgements about what actually happened (“the truth”), but stories are often appealing for other reasons \_\_because they are interesting, entertaining, familiar, reassuring, or memorable. It is well known that good stories often push out true stories and that narrative provides opportunities for innuendo, confabulation, inventing facts, suppressing facts, the use of emotive language and so on.

It is sometimes said that a story is “plausible” (“e.g. “a likely story”), if it fits into a recognizable story schema \_\_ a familiar plot of the kind Vladimir Propp analyzed or which some publishers of pulp

---

<sup>26</sup> On “cheating” see above n.000

literature provide to their authors: the love triangle; the detective gets his man; or poor girl wins prince by losing shoe. Sometimes instead of a generalized scenario the story is said to be plausible when it is analogous to a familiar (particular) story. In my experience, in preparing for mock trials, law students are suckers for such analogies: Solomon and the baby; the Prodigal Son, Cinderella, Madam Bovary, Martin Chuzzlewit; Lady Macbeth; Jack the Ripper. Their eyes light up if they find an analogy. The pleasure comes from linking the case before them to some familiar scenario or tale. It is a pleasure of recognition.

In constructing arguments about disputed facts, or in an investigation, such familiar story-types or analogies can play a useful role in imagining hypotheses and fitting the present case into a familiar pattern. But there are some obvious dangers. In advocacy they can be a seductive form of cheating. The analogy may not be exact \_\_ indeed that is highly likely. We have a glass slipper, a ball, a clock striking midnight. Can we infer \_\_ or confabulate \_\_ two ugly sisters, a prince charming, and a coach drawn by hamsters? Moreover, as with generalizations, our stock of stories are part of our so-called, allegedly shared “stock of knowledge” But notice: nearly all of my examples of stories are fictitious: they are drawn from fairy tales, myths, legends, parables, morality tales, Shakespeare plays, novels. They have little or nothing to do with empirical truth. One of the dangers of stories \_\_ especially story-types\_\_ is that they are pleasing not as aids to determining truth, but because they are entertaining, memorable, or reassuringly familiar simplifications.<sup>27</sup>

### *Generalizations*

The authors of “Sense-making” uses five broad generalizations to illustrate their account of their software. Some are standard, simplistic examples. Four are explicit: (i) “a person who takes drugs and is broke and is from an ethnic minority is likely to have bad intentions” (<11-12 >). (ii) “A witness usually tells the truth” (<5, 6 >)., (iii) “fire causes smoke” and (iv) “smoke means fire”(<p. 4 >). There is also an implicit generalization about “running away” (v) in the analysis of the *King* case (<pp. 6-7>)

(i) is an example of a prejudiced generalization. It is ruled out as unacceptable, because it violates a condition of validity, which is built into the program. With further development this could be a very

---

<sup>27</sup> Of course, where the issue before the jury is one of evaluation, for example, a judgment about reasonableness), it may be legitimate to use a fictitious parable or other morality tale. For example, in the case of *Warren (Analysis 356-78)*, one issue for the jury is whether the testator had sufficient memory of the elements of his business (here his assets and the natural objects of his bounty) to make a reasonable judgment about them. In this context, the parable of the Prodigal Son is a relevant analogy, but it can be used by both sides, (Twining (2006) at pp.397-99)

useful device for filtering out potentially prejudicial generalizations, including ones that are less blatant or clear cut.

The authors justify their use of the other simple examples “for ease of explanation”(<p.12>). It is reasonable for the purposes of exposition to non-experts to resort to simple examples and I do not object to the way these ones are used in the paper. However, “flight”, witness credibility, and suspicion (“no smoke without fire”) are useful pegs on which to hang some concerns about the kind of generalizations that are appropriate in a police investigation. So I hope that is not unfair to suggest that all five examples would be unsuitable for use in a program designed to help that kind of inquiry.

### “Flight”

In the *King* case, according to the prosecution story, after hearing a noise “Mr Z sees King running away through the closed garden door. He shouts “there is a burglar come and help me!” and runs into the garden after King. King, who wants to pretend he is lost, does not run away.” (<p.6 >) In the circumstances, the inference is reasonable because the combination of factors: noise in the house + stranger in the garden after dark = trespasser + running towards the gate = running away = fleeing.<sup>28</sup> It is going to be difficult for X to explain his presence and why he was running, because of this combination of factors can be translated into the generalization: “A trespasser in the garden fleeing under the cloak of darkness is likely to be a burglar”. Although inferring from what Mr Z saw that this was a burglar involves several inferences, in a real investigation one probably would not select this for microscopic analysis.

The inferential step from running to flight is the subject of extensive comment in the literature.<sup>29</sup> Flight, it is commonly said, implies consciousness of guilt; but running is open to several innocent explanations. If a witness hears a shout “Stop thief” and sees X, a youth, running nearby, he may wrongly infer flight, for X could have several innocent explanations: X was running *past* (e.g. jogging); X was

---

<sup>28</sup> A fuller reconstruction would be: Z hears noises in the house (a burglar?); Z looks out and sees a stranger (S) in the garden at night; G: strangers in my garden at night are probably trespassers; Z infers the stranger is a trespasser; G. Trespassers in the garden at night <+house> are often burglars/do not have good intentions; S is a burglar/ does not have good intentions; Z sees the stranger running in the garden; G. a trespasser running in the garden at night is running away; Z infers that the stranger is running away, i.e. fleeing; G: flight implies consciousness of guilt. S is conscious of guilt (about what?).

<sup>29</sup> E.g. McCormick (1999) s. 263; see *Analysis* 101, 271-2n.11 for a more detailed discussion.

running *towards* a bus stop (to catch a bus); running *after* the suspected thief; or running *for* help (e.g. running to a telephone). The shift from running to running away to flight involves a value judgment. Whether running away or flight implies “consciousness of guilt” depends on the context and the reasoner’s culture. For a middle class observer a youth running (away) from a policeman is suspicious, for someone from a different background, it may be sound common sense.

The example of “flight” illustrates several points: first, a seemingly detached description may involve a value judgment based on an unwarranted or weakly warranted inference. Second, there is not a cognitive consensus about background generalizations relating to “flight”, and third, generalizations performing the function of inferential warrants have to be *constructed* from the particular context, the available data and the reasoner’s stock of beliefs.

*“No smoke without fire”*

“Ahh, now, you see, we’ve been through this, and my thought is this: *there’s no smoke without fire.*” Archie would say, looking impressed by the wisdom of his own conclusion. “Know what I mean?” This was one of Archie’s preferred analytic tools when confronted with news stories, historical events, and the tricky day-to-day process of separating fact from fiction. *There’s no smoke without fire.* There was something so vulnerable in the way he relied on this conviction, that Samad never had the heart to disabuse him about it. Why tell an old man that there can be smoke without fire as surely as there are deep wounds that draw no blood?” (Zadie Smith (2000) at p.209)

In order to make some interesting points about the relationship between reasoning with causal (fire causes smoke) and evidential (smoke means fire) generalizations, “Sense-making” postulates that smoke is usually caused by fire but that in exceptional cases it is caused by a smoke machine \_\_ i.e. it is defeasible. I do not want to question the argument in this context, but this example illustrates some further points about defeasible generalizations in argumentation in legal contexts.

This is another example of a generalization expressed in very abstract terms without any reference to context. The generalization makes more sense in the bush of New South Wales than in smoke-filled rooms on Capitol Hill. It is much too broad for us to be able to be confident about how frequently the generalization holds. Even if we are confident about what the generalization means (see

below), the question whether this is a generalization that is sufficiently strong to ground a presumption that it applies unless challenged, is unknown and probably unknowable.<sup>30</sup>

Second, it is unclear what count as “smoke” and “fire”. Is a cigar on fire? Is a match lighting a cigar on fire, but not the cigar? Are exhaust fumes an example of smoke without fire? Or steam from a kettle? We could get some scientific definitions of both terms and some well-founded generalizations,<sup>31</sup> but would these be appropriate in the context of criminal investigation? If we are relying on an ordinary witness, are we not more concerned with smoke-like phenomena, which might include steam, fumes or other vapours, or in some contexts, mist or fog? If W1 testifies “I saw smoke”, meaning I saw something that resembled smoke, are we justified in inferring “probably there was a fire” because “smoke means fire”?

Third, we have so far taken the statement literally. But there is a metaphorical usage more germane to criminal investigation. “When there are suspicious circumstances, there has possibly/probably been a crime”. I own a book by a former police officer called *The Signs of Crime* (Powis, 1977) in which the author collected together dozens of examples of police folk wisdom about grounds for suspicion.<sup>32</sup> Many take the form of what Sherlock Holmes called “trifles”. Some are quite illuminating. For example: if, on a fine day, one sees a car being driven with its windscreen wipers going, one may infer that the driver is not familiar with the car; if someone is driving a car with which they are not familiar, it may have been stolen. (Powis, 1977 at p.6) So this is a suspicious indicator worth further investigation. Of course, as the author concedes, there are other possible explanations. The driver is driving a rented car for the first time; the windscreen was dirty; the driver is an absent-minded professor; the mechanism is faulty; and so on. Nevertheless this is “a sign of crime” — a ground for suspicion. What is the probability that

---

<sup>30</sup> The idea of defeasible evidential generalizations as “prima facie reasons” is a useful one, but it should not be confused with the legal concept of presumptions. (Gaskins, 1992) A prima facie reason is a convenient simplifying device, but it can be dangerous if it is formalized or treated as more than that. An assertion that a particular generalization is a prima facie reason needs to be subject to critical scrutiny (for a suggested protocol for assessing the plausibility and validity of a generalization in the context of an argument, see *Analysis* 279-80).

In circumstances X, smoke without fire is exceptional; but in circumstances Y, smoke without fire is the norm.

<sup>31</sup> Such generalizations would tend to be much more specific, taking the form: “*In circumstances X, Y frequently etc* ..... [Note on EBM and analogies> Random clinical trials usually use “normal or “average” patients between 20 and fifty years old The findings are often however extended to children and old people.]

<sup>32</sup> David Powis, *The Signs of Crime: A Field Manual for Police* (New York: The John Jay Press (1977). The author was Deputy Assistant Commissioner, Metropolitan Police in London but the book was published in the United States.

the car has in fact been stolen? This depends on the relative frequency of the various explanations. But often we have no basis for making such estimates of frequency.<sup>33</sup>

Many of the other generalizations in *Signs of Crime* are not so innocent — they are more like a catalogue of prejudices about teen-agers, women, ethnic minorities, short people, foreigners, homosexuals and so on. The book can be read as a compendium of ageist, sexist, racist, anti-semitic, xenophobic, homophobic prejudices. Almost anything abnormal is a sign of crime. Common sense generalizations are frequent vehicles of prejudice. This is not the kind of thinking one wishes to encourage in policemen, let alone factor into into machines. But *Signs of Crime* illustrates the kind of generalizations police investigators use in practice. Many of the experience-based generalizations cited in the book contain genuine insights. The ability to make sensitive differentiations between well-based factual generalizations and prejudicial ones and how to act when they overlap is, one hopes, part of police training. But one also hopes that police software will reinforce rather than undermine such training. Stories and generalizations, are both necessary and very dangerous. Investigators need to be warned of these dangers. But how to build this into software program intended as an aid to reasoning is inherently problematic.

Perhaps even more worrying, in this context, is how the author of this book slips easily from “worth noticing”, to “worth investigating”, to “suspicious” to “highly suspicious”. In the example of seeing windscreen wipers being on a fair day, we have hardly any basis for estimating the relative frequency of the various explanations. But if one has to speculate, one might say that this indicator does

---

<sup>33</sup> The whole passage reads:

**Watch for**

- Persons, especially those in the suspect age group, who appear unfamiliar in any way with the mechanism or controls of the vehicle they are using. Switching lights on in daylight or windscreen wipers on in dry weather, without realizing the switch has been used, is very suspicious indeed. In some vehicles inexpert shorting of the ignition to start the vehicle without a key can cause the windscreen wipers to operate and, more important refuse to stop! In dry weather, windscreen wipers operating for more than a few seconds are *always* worth prompt investigation”. (id. at p.6).

Note that the simple generalization “windscreen wipers on in dry weather is suspicious” is qualified by several factors, not all of which will be observable at first sight: age of the driver; wipers operating for more than a few seconds; car started without a key. The most easily observable fact of wipers working in dry weather only suggests that it is slightly more likely that this car may have been stolen than other similar cars. [Or, to be more exact, since it is *possible* that any car we see has been stolen, this indicator slightly raises the likelihood.] Would a stop and search be justified if the wipers were continuously on and the driver is a youth?

no more than suggest that it there is a chance that the car has been stolen.<sup>34</sup> So which is the defeasible generalization and which the exception? Is it: “If a car has its windscreen wipers going for more than a few seconds, the most likely explanations are innocent, but exceptionally it may have been stolen” or is it: “When a car has its windscreen wipers going for more than a few seconds, it probably has been stolen, but exceptionally there may be an innocent explanation”? What is worrying is that the latter judgement (“very suspicious”, “*always* worth investigation”) was in a book intended as a training manual for police officers on both sides of the Atlantic.

Some of the obvious overt prejudices displayed in the book would not be as acceptable today as they may have been in 1977. To that extent this illustrates how “common sense” and our “stock of knowledge” vary over time and context. But others, particularly less blatant ones, illustrate why in analysis of evidence generalizations need to be treated with great caution and those that play a significant part in an argument need to be articulated with care and critically assessed.

### *Conclusion*

The underlying theory of the software program described in “Sense-making” complements modified Wigmorean analysis (MWA). Both adopt a qualitative rather than a quantitative approach. (<p. 2-3 >) MWA is broadly compatible with the kind of logic involved, including abductive inference to the best explanation and the idea of defeasible argumentation. Both approaches are mainly valuable as aids to thinking, especially constructing and evaluating arguments, rather than as methods of presenting them in order to persuade. Both approaches can be applied at different stages of criminal investigation (and more broadly of legal processes), but the specific device of Wigmore charts (one part of MWA) is more useful in hypothesis-testing and discarding than in hypothesis formation, which typically requires imaginative reasoning (*Analysis* (2005) Ch. 2). Anchored narratives (AN) and MWA have similar theoretical assumptions, except that MWA gives a radically different account about the

---

<sup>34</sup> Of course, it is possible that any vehicle noticed by a police officer has been stolen. The fact that the wipers are going on a dry day slightly increases the likelihood, but that may still be well below 50%. The sensible next step would be to ask some more preliminary questions before taking action: is the vehicle being driven normally? Is it of a type that attracts local thieves? (local experience-based knowledge); does the driver look like a car thief? (experiential “knowledge” verging on dubious practices of profiling).

relationship between stories, generalizations and argument. The proposed program has considerable promise, but before it can be of positive practical value in police investigation, more attention needs to be given not only to the obvious dangers of using stories and generalizations in this context, but also about what positive guidance can be given to mitigate these dangers. There is, however, an unresolved tension between the simplifying tendencies of formalized computer programming and the tendency of MWA to emphasize the complexities of practical inferential reasoning and argumentation in legal contexts.