2

Evolving Datacase Designs

Robert W. Lawler

Computer-based tools and techniques can be used to advance the state of the art in case study analysis so that this "would-be science" can cross the frontier of credibility separating fields of study and the sciences.

This chapter describes in summary a set of designs for databases of case study materials (call them *datacases*) and analyzes both their potential value and some limitations encountered in trial implementations. These datacase designs were developed over a period of 12 years and partially implemented on various computer systems. The first was begun in Xerox's Notecards, with later implementations in different versions of Macintosh Hypercards. This discussion is essentially an exploration-based user-requirements study for a facility to support case study analysis through the comparison of databases. The final design discussion looks toward cross-platform portability as an aspect of designs for the future; such is essential because of the long, time and labor-intensive characters of case study analysis.

The text is organized around the centrality of the analyst-user's need to make sense of a variety of information in various online files interconnected in different ways. That second aspect of the designs may make the ideas advanced here useful to others facing similar challenges. To supplement the primary text, this chapter includes figures, appendices, and cameos. The latter are short descriptions of ideas and situations or arguments that are complete in themselves. They are off the main line of the primary text, but all are essential in illuminating the meaning and supporting the argument of the primary text. Because the cameo texts were written at various times, they are "time stamped" to more precisely specify their relation to the primary text.

HYPERTEXT FOR CASE STUDY ANALYSIS

A few years ago Hypertext[1] seemed an idea whose time had come. At the First National Hypertext Workshop,[2] there was enthusiasm, new products coming online (Apple's Hypercard was then a novelty leading the way), and a range of systems representing broad marketing possibilities. The flood tide of enthusiasm is past; some commercial enterprises are decommiting production of hypertext systems, and voices are heard with the complaint that creating hypertexts is hard.[3] Even enthusiasts must admit that there are a few, if any, stunningly effective uses of hypertext that clearly justify the effort of their construction.[4] Contrast this situation with the introduction on personal computers of word processing or spreadsheets.

A critic might even complain that, despite commendable efforts, there is yet not any broadly successful, serious work where hypertext is the native mode of the work itself. The situation today is that no one knows yet what it would mean to use hypertext with artistry. Where is the Shakespeare or the Homer of hypertext? The artistry of composition in hypertext will be one of some different sort, one that is still to be discovered and invented. Making a contribution to that effort is one of the objectives here.

I believe that hypertext as a form is still promising and that some promises are credible, but a deeper penetration is needed to develop the potential of

---

[1] According to enthusiasts, text as we know it today will be superseded by hypertext, a machine-resident body of information designed to take advantage of new forms of organization and interrelation that electronic media make possible. People have been working on hypertext in laboratories for years. An influential system, Xerox Notecards, was developed by Frank Halasz (1985) and colleagues at PARC (the Palo Alto Research Center of the Xerox Corporation). With the Notecards system metaphor, text and graphics are developed on cards that are stored in file boxes. The metaphor is then modified to permit hierarchical organization, for file boxes may be stored in file boxes ad infinitum.

[2] The primary papers of this conference are presented in a special issue of *Communications of the ACM*. That contains two papers of special importance, reflecting these two talks. As part of a retrospective and specification of new directions, Halasz set out an agenda of seven issues to be addressed in future research and implementation efforts. In his keynote speech at that same conference, Andries Van Dam, leader of another early and primary center for hypertext research at Brown University, presented his list of next generation enhancements.

[3] In a letter to *Technology Review Magazine* (MIT, November–December, 1990) a multimedia developer argued that the production of hypertext could be compared to the creation of a film based on a novel, and that the layering of additional requirements on top of an already formidable task implied the scope of effort required to produce a good example of the kind would be enormous.

[4] Several years ago, McClintock (1986) argued that computing technology had not had a significant effect on higher education in many fields for a specific reason. Even though there was no shortage of demonstrations, he argued, there was a paucity of material for serious study. Going beyond tool making to the production of useful materials, major works that are capable of exploiting the new opportunities of technology is a frontier challenge for those who work at technology in education.

hypertext in concrete uses. This is a time to go beyond enthusiasm, a time to ask what the essential value of hypertext is and whether, on the one hand, it can be of profound importance in achieving ends that are of obvious value or, on the other hand, permit achieving objectives that were not imaginable previously. An approach to the issue of new objectives can be made through the way that theory and data relate in case study analyses.

Relating Theory and Data

The commitment to examination of extensive evidence may bring case study into contempt as an "empirical method." There is certainly a more intimate entangling of theory and data in case study analyses than in most other methods. We can describe this situation with some well-known Piagetian terminology. One assimilates data to a theory when possible, and one accommodates theories to a datum when necessary. This sort of interaction is not an inferior way of doing science, but it is the usual way in which inferences are made. In support of this claim I cite the comments of Goodman (1983) from his discussion of inference in *Fact Fiction, and Forecast* (see Cameo 1 for his text, then Figure 2.1 for a graphic summary).

CAMEO 1. INFERENCE—AS ASSIMILATION AND ACCOMMODATION

The validity of a deduction depends not upon conformity to any purely arbitrary rules we may contrive, but upon conformity to valid rules. When we speak of the rules of inference we mean the valid rules—or better, some valid rules, since there may be alternative sets of equally valid rules. But how is the validity of rules to be determined? Here again we encounter philosophers who insist that these rules follow from some self-evident axiom, and others who try to show that the rules are grounded in the very nature of the human mind. I think the answer lies much nearer the surface. Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we actually make and sanction. If a rule yields unacceptable inferences, we drop it as invalid. Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences.

This looks flagrantly circular. I have said that deductive inferences are justified by their conformity to valid general rules, and that general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. A rule is amended if it yields an infer-

The page make-up of all cameos & figures needs to be improved in a uniform fashion.
RWL.

ence we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either.

All this applies equally well to induction. An inductive inference, too, is justified by conformity to general rules, and a general rule by conformity to accepted inductive inferences. Predictions are justified if they conform to valid canons of induction; and the canons are valid if they accurately codify accepted inductive practice. (Goodman, 1983, pp. 63–64)

Further, if we are willing to recognize that case analysts function in a mode of theory development, as opposed to a mode of theory testing, that recognition is not merely acceptable, it is important. For then it is no surprise that case studies have been productive of the most important social science theories of the

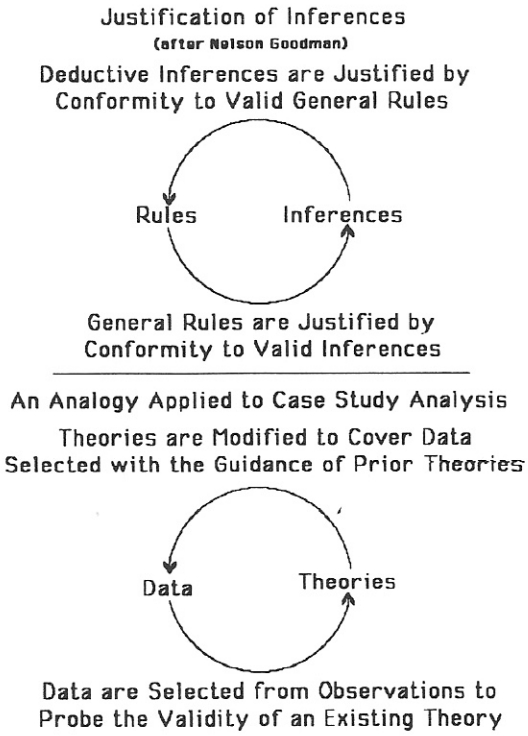


Figure 2.1. An analogy, after Goodman on inference.

century. It argues further that case study analysis is as close as we can come, in the social sciences, to that kind of problem solving behavior through which scientists learn the implications of the theories they hold and how to evaluate applicability of theories to problematic situations.[5]

My Work With Case Analysis Before Hypertext

I can believe in the difficulty of constructing usable, valuable hypertexts because it has been my continual effort for more than a dozen years, in service of a scientific goal to which my commitments go back an even longer time. My special focus has been on case study as a method of exploring learning. The long-time scale and labor-intensive character of case study emphasizes specific considerations when developing tools to support analytical work.

Working with hypertext involves both technical and conceptual difficulties at the same time. I believe the nature and value of linking have been insufficiently deeply understood—at least with respect to the kinds of applications examined in this chapter. In a subsequent section I advance the notion that we need to implement within hypertext systems a new kind of link, one whose existence is so important to the use of hypertexts that it would be fair to describe it as "the missing link" whose lack has held back the general, effective, and powerful use of hypertext.

Let us look at some of the circumstances of doing similar work before the availability of computer-based word processing and hypertext. I completed a thesis reporting my doctoral research in 1977–1979 (Lawler, 1979a). The project analyzed in detail the development of a single child over a 6-month period in a variety of domains. The thesis was 625 pages.[6] The materials supporting the observations selected for that document were much more extensive. I look now at my bookshelves and see three large volumes of manual transcriptions of tape recordings, an equally large collection of "vignettes"—short stories written about the child's development during the course of the period of intense observation,—a smaller collection of descriptions of experiments from the Piagetian repertoire undertaken with the subject, a large notebook of early papers and preliminary drafts related to the same corpus, and—untranscribed—more than a dozen notebooks used as journals during the period.

The effort was undertaken not as a theory-driven hypothesis-disconfirmation experiment, or merely as an empirical exercise. The primary notion I had in mind was to develop a sufficiently concrete collection of empirical material that would permit me the opportunity to provide a worked example of development

---

[5] So Lewin (1935) argued in his famous paper, "On the Conflict Between Aristotelian and Galilean Modes of thought in Contemporary Psychology."

[6] The text was double-spaced, but 10 point.

using computing-oriented descriptions of knowledge and its changes.[7] The thesis, *One Child's Learning* (Lawler 1979b), was an integration of *ideas* derived from theoretical concerns and an *interpretation*. The latter was constructed from selected incidents in the corpus of these documents and 80 half-hour videotapes; they were reviewed over a multiyear period of analysis and interpretation. Obviously the thesis did not contain all the information in the corpus; nor, later, did the books based in part on the thesis contain all of its content—although both the corpus and the books were extended by other materials. As new material, notions, and interpretations were added, the control and integration of the ideas became progressively more difficult, until I found myself repeating too often the lament of Emily Dickinson:

> The thought behind I strove to join
> Unto the thought before
> Till sequence raveled out of reach
> Like balls upon the floor.

When this material was prepared for academic articles, it was, of course, rewritten. A 30-page journal article cannot well contain extensive selections from the corpus of a long developmental study.[8] Ultimately, I decided to not compress all the material of the corpus for journal articles but rather to prepare a coherent selection of material and analyses as a book, because the process of rendering the material accessible also reduced its credibility. My hope is that in the future we can use hypertext to make case studies more accessible without reducing their credibility.

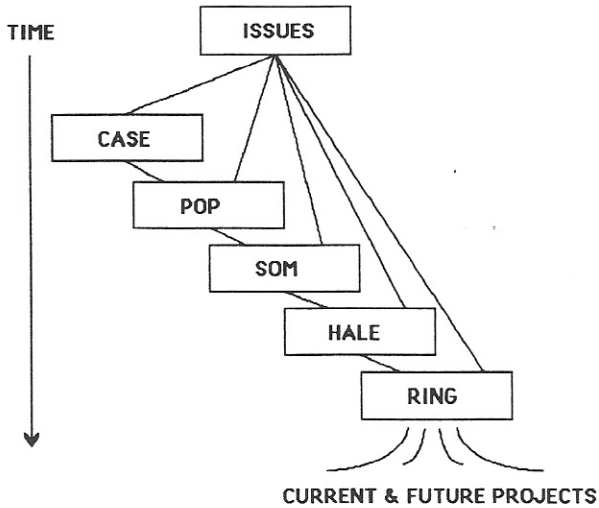
Being Overwhelmed by Material

When I observed the effort that had been expended in completing the analysis of *One Child's Learning* as a thesis—and looked forward to what I foresaw in bringing my work to a broader public as a book—I knew at once that I needed all the help I could hope for from computing.[9] The threat of being overwhelmed by a morass of data—a common fate of failed case studies—was even stronger because I had begun constructing a second major case study corpus during the

[7] People have criticized Piaget because he never provided a worked example of how his theory could be applied in a specific case to a particular child developing significant new knowledge. That was not his intention, so criticizing him for that reason is not fair. Such was, however, my ambition, and my work is vulnerable to criticism to the extent that it fails to do so.

[8] The first analytical paper from this study, Lawler (1981), can represent the compression appropriate for an academic article. In a summary article, such as Lawler (1989), research-based arguments become pretty notions decorated with hopefully memorable detail.

[9] Because I had spent 8 years as database and telecommunications applications engineer for IBM before starting my doctoral program, I knew that computing could be a significant help when systems were properly designed.



CASE - A Case Analysis Support Environment, prototype for a theory development facilty
POP - The Psychology of the Particular, a datacase based on analyses and a corpus created by Robert W. Lawler
SOM - The Society of Mind, a version in Hypercard based on materials provided by M. Minsky
HALE - Hartley Hale, a datacase based on analyses and a corpus created by Robert W. White
RING - A design for Datacase Link Authoring Facility

Figure 2.2. A guide for the reader.

same period of time. My initial attempt to cope with that second corpus is the subject of the next section of this chapter. But before plunging into the detail of that project, let me pause here to offer a little guidance to the reader on what is found in the remainder of this chapter.

A Guide for the Reader

Figure 2.2 represents both the sequence of four projects in time and their order of presentation here. Thus it is not surprising that each one affected the successors. Because they all involve the interrelationships of theories, data, and systems of hypertexts, there are also a series of common issues generating the various projects. The internal structure of each discussion is generally the same, introducing the motivation for the work, a synopsis of what was done, and a summary of the lessons learned from the experience. Each is represented by an

acronym, spelled out in the text below Figure 2.2. Following the exposition and discussion of this work, the text proceeds to discuss my best current design for a Case Analysis Support Environment (CASE), embodied in a software shell I call RING. The conclusion then discusses the general usefulness of this environment for other kinds of text-based analysis.

A CASE ANALYSIS SUPPORT ENVIRONMENT

Motivation for the Work—A Technical Vision

It took me 10 years to bring my doctoral study to a traditional conclusion, one that can be represented by the publication of a coherent interpretation in books (Lawler 1985; Lawler, duBoulay, Hughes, & Macleod, 1986). Despite its commitment to empirical material, the study was theory driven (see Appendix A).[10]

During that same period of time, I began collecting materials for the corpus of a second case study. Why, while still overwhelmed by the task of making sense of one study, take on a second? Opportunity might be one reason, given that a child was born to our family then and the second corpus was, obviously, an infant study. More importantly, I saw there an opportunity to address a real issue and move it from the arena of ideological dispute into the study—although not the laboratory. That issue was the character of language learning and its relatedness or lack of relation to other kinds of learning.[11]

My second case study corpus was organized to track both the infant's language and sensorimotor learning in the social context of the home. The amount of material collected and inclusion of an extensive videotaped component urged on me the need to develop some computer-based tools to help with the administration and interpretation of such a large-scale case study corpora. I discussed this issue with Marvin Minsky at that time (Lawler, 1981), showing him the diagram in Figure 2.3, as my sketch for the kind of facility that could help support the analyst in an attempt to develop theories related to extensive areas of evidence. Minsky suggested I take seriously the notions of Nelson and consider "hypertext" as a candidate form of database organization for such a facility. I

---

[10] It was my conviction then—and it remains so now—that the best work in the artificial intelligence community at that time, however lucid and elegant, was too remote from the details of everyday life to bring computational theories into confrontation with the complexity of the problems they tried to address. For contrast, compare Winston (1975) and Sussman (1975) with Barker and Wright (1955/1971). The question of how any computational machine (such as a human mind is) could cope with relations between the messy richness of details in experience and possible complexities of representation was one of the issues generating Minsky's (1975) well-known paper "A Framework for Representing Knowledge."

[11] That issue was somewhat contentious at the time. Witness the discussions between Piaget and Chomsky and their colleagues at the Royaumont Abbey. Some aspects of that debate are preserved in the book *Language and Learning* (Piatelli-Palmerini, 1980). The issues are not resolved.