Detecting United States Mediation Styles in the Middle East, 1979-1998

Philip A. Schrodt
Department of Political Science
University of Kansas
Lawrence, KS 66045 USA
phone: +1.785.864.9024 fax: +1.785.864.5700
p-schrodt@ukans.edu

April 1999

Paper presented at the International Studies Association meetings
Washington, DC

The event data and computer programs used in this analysis, as well as a PDF version of this paper, can be downloaded from the Kansas Event Data System project web site:
`http://www.ukans.edu/~keds`

My thanks to Ze'ev Maoz, Alex Mintz, Cliff Morgan, Glenn Palmer and Richard Stoll and their research institutes for their organization and sponsorship of the "Multiple Paths to Knowledge Project," which motivated this research.

ABSTRACT

This research is part of the "Multiple Paths to Knowledge Project" sponsored by the James A. Baker III Institute for Public Policy, Rice University, and the Program in Foreign Policy Decision Making, Texas A&M University. The paper deals with the problem of determining whether the mediation styles used by four U.S. Secretaries of State—George Schultz, James Baker, Warren Christopher and Madeline Albright—are sufficiently distinct that they can be detected in event data. The mediation domain is the Israel-Palestinian conflict from April 1979 to December 1998, the event data are coded from the Reuters news service reports using the WEIS event coding scheme, and the classification technique is hidden Markov models.

The models are estimated for each of the four Secretaries based on 16 randomly chosen 32-events sequences of USA>ISR and USA>PAL events during the term of the Secretary. Each month in the data set is then assigned to one of the four Secretarial styles based on the best-fitting model. The models differentiate the mediation styles quite distinctly and this method of detecting styles yields quite different results when applied to ISR-PAL data or random data. The "Baker" and "Albright" styles are most distinctive; the "Schultz" style is least; both results are consistent with many qualitative characterizations of these periods.

A series of t-tests is then done on Goldstein-scaled scores to determine whether the mediation styles translate into statistically distinct interactions in the ISR>USA, ISR>PAL, PAL>USA and PAL>ISR dyads. While there are a number of statistically-significant differences when the full sample is used, these may be due simply to the overall changes Israel-Palestinian relations over the course of the time series. When tests are done on months that are out-of-term—in other words, where the style of one Secretary is being employed during the term of another—few statistically-significant differences are found, though there is some indication of a lag of a month or so between the change in style and the behavioral response. It appears that the effects of the differing styles are not captured by changes in aggregated data, possibly because these scales force behavior into a single conflict-cooperation dimension.

(In keeping with the protocols of the "Multiple Paths to Knowledge" conference, I have placed into the text of the paper an assortment of pedagogical notes concerning the process of constructing it.)

This paper grew out of a collaboration that came together during the Rice meeting. Jonathan Wilkenfeld (Maryland), Charles Taber (SUNY Stony Brook) and I were assigned to a group, and realized that our research methods—experimental methods, artificial intelligence and event data respectively—presented a natural hierarchy of control. I could look at the problem of mediation using event data covering a real-world problem: with this approach, I had no controls on the data, and simply had to take what the historical record gave me. Wilkenfeld could use a formal experimental design, where he could control some aspects of the situation (notably the tactics used by the mediator and the presentation of a problem, in his case one abstracted from a Canada-Spain fisheries dispute) and measure others (notably the "cognitive complexity" of his subjects) but he could not actually find out why the subjects were behaving as they did. Finally, Taber's artificial intelligence simulations would give him complete control of every aspect of the situation, including the problem-solving methods used by the simulated actors.

The three of us were generally acquainted with each other's work—and Taber and I had been on a number of panels together over the years—but we had never actually worked together. We did, however, share a common dissertation advisor—Dina Zinnes—which might have contributed to a common outlook on how one should do research. The coordination between the projects was very loose, with occasional exchanges of email and a few decisions on common vocabulary.

This research is part of the "Multiple Paths to Knowledge Project" sponsored by the James A. Baker III Institute for Public Policy, Rice University, and the Program in Foreign Policy Decision Making, Texas A&M University. This paper deals with the problem of determining whether the mediation styles used by four U.S. Secretaries of State—George Schultz, James Baker, Warren Christopher and Madeline Albright—are sufficiently distinct that they can be detected in event data, and whether the different mediation styles have different results on the course of the conflict. The mediation domain is the Israel-Palestinian conflict from April 1979 to December 1998, the event data are coded from the Reuters news service reports using the WEIS coding system, and the modeling technique is hidden Markov models.

Why this data set and model? Simple: the data were readily available and I figured the model would work. I had been working with the Levant data set for a number of years—as well as spending considerable time in the region itself—and was both comfortable with the data and confident of its quality. The Israeli-Palestinian conflict is highly conspicuous and most U.S. Secretaries of State try to put their individual stamp on mediating the conflict, so I anticipated that I could find distinctive styles.

The original paper proposal, reflected in the title appearing in the International Studies Associations program, had a parallel analysis with data from the on-going crisis in the former Yugoslavia. I encountered some problems obtaining this data and wasn't able to incorporate it, which weakened the overall analysis.

Since 1967, the Israeli-Palestinian dispute has been one of the key foreign policy foci of the United States. Despite this region being almost totally insignificant on most traditional measures geopolitical importance—its total population and area are smaller than that of many metropolitan areas, and it controls no vital trade routes or resources[1]—its symbolic importance places the conflict high on the priority list of U.S. Secretaries of State. Warren Christopher, for example, made some twenty major trips to the Middle East and only one to China. From the "shuttle diplomacy" of Henry Kissinger to the Camp David Accords of Jimmy Carter through the recently-concluded 1998 Wye River Agreement, Middle East diplomacy has played a major role in virtually every administration over the past 30 years.

The assumption of this analysis is that the various U.S. Secretaries of State have had distinctive styles in attempting to mediate the Israeli-Palestinian dispute. At a qualitative level, this assumption is relatively uncontroversial (see for example Gerner 1991; Quandt 1993; Tessler 1994; Eisenberg & Caplan 1998; and Guyatt 1998). The purpose of this paper is to ascertain whether those styles can be systematically characterized using international event data and what, if any, difference these styles make in the resolution of the conflict.

Hidden Markov models[2]

Techniques for analyzing sequences of discrete events—nominal-level variables occurring over time—are poorly developed compared to the huge literature involving the study of interval-level time series. Nonetheless, several methods are available, and the problem has received considerable attention in the past three decades because it is important in the problems of

---

[1] On the eve of the first elections to the Palestinian National Council in January 1995, Albert Aghazarian, the public-relations director of Birzeit University, began his remarks—to an East Jerusalem hotel ballroom crowded with the international media—by observing, "Let me say that I am pleased that so many of you have travelled thousands of miles and spent many tens of thousands of dollars to observe what is essentially a municipal election..."

[2] This description is shamelessly lifted from Schrodt (1999) with modifications.

studying genetic sequences in DNA, and computer applications in involving human speech recognition. (Both of these problems have potentially large economic payoffs, which tends to correlate with the expenditure of research efforts.)

Hidden Markov models (HMM) are a recently developed technique that is now widely used in the classification of noisy sequences into a set of discrete categories (or, equivalently, computing the probability that a given sequence was generated by a known general model). A sequence is "noisy" when it contains missing, erroneous and extraneous elements, and consequently the sequence cannot be classified by simply matching it to a set of known "correct" sequences. A spell checking program, for example, would always mark "wan" as an incorrect spelling of "one", because written English usually allows one and only one correct spelling of a word. Spoken English, in contrast, allows a wide variation of pronunciations, and in some regional dialects, "wan" is the most common pronunciation of "one". A computer program attempting to decipher spoken English needs to provide for a variety of different ways that a word might be pronounced, whereas a spelling checker needs only to know one.

While the most common applications of HMMs are found in speech recognition and comparing protein sequences, a recent search of the World Wide Web found applications in fields as divergent as modeling the control of cellular phone networks, computer recognition of American Sign Language and—inevitably—the timing of trading in financial markets. The standard reference on HMMs is Rabiner (1989), which contains a thorough discussion of the estimation techniques used with the models as well as a notation that is used in virtually all contemporary articles on the subject.

> *The hidden Markov model, like the data, was readily available. However, getting to that model had been a long process with numerous false starts. I had been searching since about 1984 (when I began working with computational modeling and artificial intelligence techniques) for a reliable method of characterizing international event sequences so that I could use these in a form of "reasoning by analogy." There is no comparable problem in econometrics and finding an appropriate method proved fairly difficult—at various points I experimented with Levenshtein metrics (Schrodt 1984, 1991), genetic algorithms (Schrodt 1986, 1989), and some customized techniques for decomposing the sequences (Bennett and*

Schrodt 1987, Schrodt 1990).[3] *After a long hiatus from this issue due to demands from the Kansas Event Data System project and early warning models, I finally returned to it when I read about at article on HMMs in* The Economist *magazine in 1997. I tracked down some code on the World Wide Web, started experimenting, and have been working with the method ever since, albeit primarily in the context of early warning.*

An HMM is a variation on the well-known Markov chain model, one of the most widely studied stochastic models of discrete events (Bartholomew 1975). Like a conventional Markov chain, a HMM consists of a set of n discrete states and an n x n matrix [A] = $\{a_{ij}\}$ of *transition probabilities* for going between those states. In addition, however, every state has a vector of *observed symbol probabilities* that combine into a second matrix [B] = $\{\mathbf{b}_j(k)\}$ corresponding to the probability that the system will produce a symbol of type k when it is in state j. The states of the HMM cannot be directly observed and can only be inferred from the observed symbols, hence the adjective "hidden". This is in contrast to most applications of Markov models in international politics where the states correspond directly to observable behaviors.

While the theory of HMM allows any type of transition matrix, the model that I will be testing allows transitions only to the previous state and the next state (as well as remaining in the current state). This is an extension of the unidirectional "left-right" (LR) model that is widely used in speech recognition. The transition matrix **A** is therefore of the form

$$\begin{pmatrix} a_{11} & 1-a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \dots & 0 \\ 0 & a_{32} & a_{33} & a_{34} & \dots & 0 \\ \dots & & & & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_{n-1,n} \\ 0 & 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

and the individual elements of the model look like those in Figure 1. I will refer to this as a "left-right-left" (LRL) model; a series of these individual elements form an HMM such as the 6-state model illustrated in Figure 2. .

[3]This paper was written when D. Scott Bennett was an undergraduate at Northwestern; he would appear at the "Multiple Paths to Knowledge" conference as an associate professor at Penn State.

Figure 1. An element of a left-right-left hidden Markov model

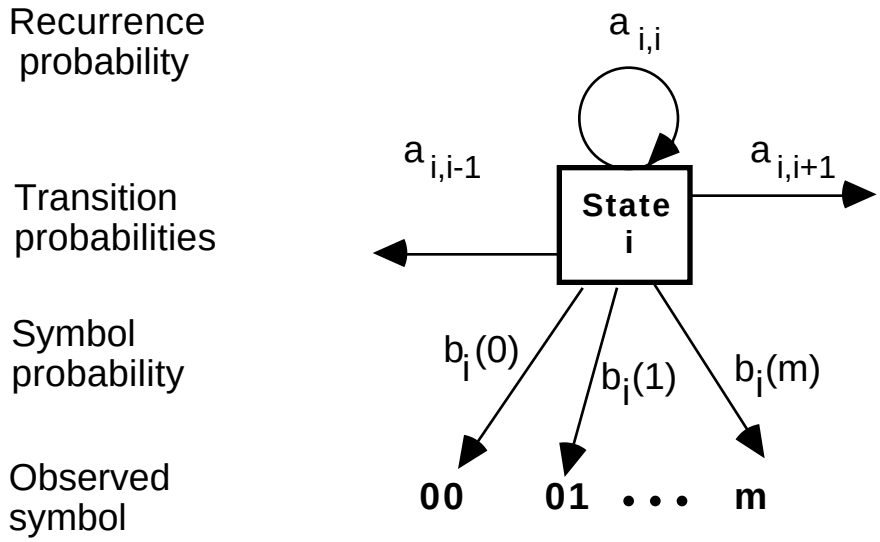
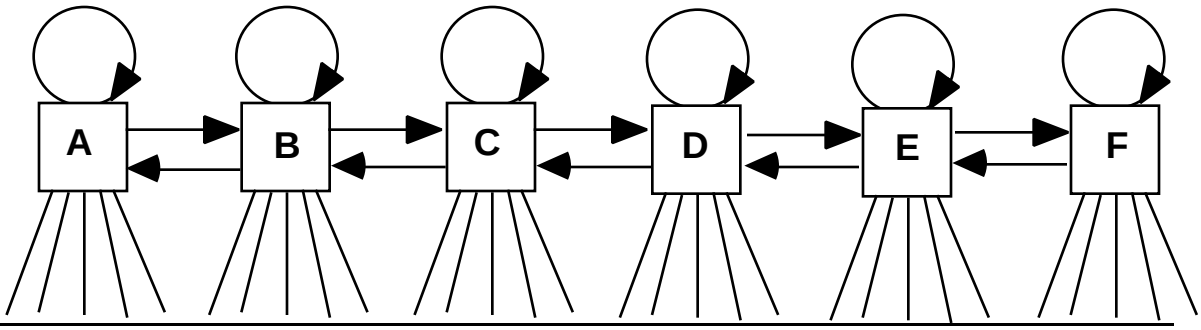


Figure 2. A left-right-left (LRL) hidden Markov Model



In empirical applications, the transition matrix and symbol probabilities of an HMM are estimated using an iterative maximum likelihood technique called the Baum-Welch algorithm.[4] This procedure takes a set of observed sequences (for example the word "seven" as pronounced by twenty different speakers, or a set of dyadic interactions from an event data set) and finds values for the matrices **A** and **B** that locally maximize the probability of observing those sequences. The Baum-Welch algorithm is a nonlinear numerical technique and Rabiner

[4] Rabiner (pg. 253) notes that the Baum-Welch algorithm is equivalent to the more familiar "expectation-modification" (EM) approach of Dempster, Laird and Rubin.

(1989:265) notes "the algorithm leads to a local maxima only and, in most problems of interest, the optimization surface is very complex and has many local maxima."

Once a set of models has been estimated, it can be used to classify an unknown sequence by computing the maximum probability that each of the models generated the observed sequence. Once the probability of the sequence matching each of the models is known, the model with the highest probability is chosen as that which best represents the sequence.

For example, in a typical speech-recognition application such as the recognition of bank account numbers, a system would have HMMs for the numerals "zero" through "nine". When a speaker pronounces a single digit, the system converts this into a set of discrete sound categories (typically based on frequency), then computes the probability of that sequence being generated by each of the ten HMMs corresponding to the ten digits. The HMM that has the highest probability—for example the HMM corresponding to the numeral "three"—gives the best estimate of the number that was spoken.

The application of the HMM to the problem of generalizing the characteristics of international event sequences is straightforward. The symbol set consists of the event codes taken from an event data set such as WEIS (McClelland 1976). The states of the model are unobserved, but have a close theoretical analog in the concept of crisis "phase" that has been explicitly coded in data sets such as the Butterworth international dispute resolution data set (Butterworth 1976), CASCON (Bloomfield & Moulton 1989, 1997) and SHERFACS (Sherman & Neack 1993), and in work on preventive diplomacy such as Lund (1996). For example, Lund (1996:38-39) outlines a series of crisis phases ranging from "durable peace" to "war" and emphasizes the importance of an "unstable peace" phase. In the HMM, these different phases would be distinguished by different distributions of observed WEIS events found in the estimated $\mathbf{b_j}$ vectors. A "stable peace" would have a preponderance of cooperative events in the WEIS **01-10** range; the escalation phase of the crisis would be characterized by events in the **11-17** range (accusations, protests, denials, and threats), and a phase of active hostilities would show events in the **18-22** range. The length of time that a crisis spends in a particular phase would be

proportional to the magnitude of the recurrence probability a_{ii}. This approach is easily generalized to mediation—a mediator has a series of strategies (e.g. talking, threatening, rewarding) that he or she follows over time, and moves between these—for example doing a "good cop/bad cop" routine—depending on circumstances.

The HMM has several advantages over alternative models for sequence comparison. First, if N<<M, the structure of the model is relatively simple. For example a left-right model with N states and M symbols has 2(N-1) + N*M parameters compared to the M(M+2) parameters of a Levenshtein metric, another commonly used sequence comparison method (see Kruskal 1983; Sankoff & Kruskall 1983). HMMs can be estimated very quickly, in contrast to neural networks and genetic algorithms. While the resulting matrices are only a local solution—there is no guarantee that a matrix computed by the Baum-Welch algorithm from a different random starting point might be quite different—local maximization is also true of most other techniques for analyzing sequences. Furthermore, the computational efficiency of the Baum-Welch algorithm allows estimates to be made from a number of different starting points. The HMM model, being stochastic rather than deterministic, is specifically designed to deal with noisy input and with indeterminate time (see Allan 1980); both of these are present in international event sequences.

HMMs are *trained by example*—model that characterizes a set of sequences can be constructed without reference to the underlying rules used to code those sequences. This provides a close parallel to the method by which human analysts generalize sequences: They typically learn general characteristics from a set of archetypal cases.

HMMs do not require the use of interval-level scales such as those proposed by Azar and Sloan (1975), Vincent (1979) or Goldstein (1992). These scales, while of considerable utility, assign weights to individual events in isolation and make no distinction, for example, between an accusation that follows a violent event and an accusation during a meeting. The HMM, in contrast, uses only the original, disaggregated events and models the context of events by using different symbol observation probabilities in different states. An event that has a low probability within a particular

context (that is, a specific hidden state) lowers the overall probability of the model generating the sequence. In aggregative scaling methods, events have the same weight in all contexts.

While most existing work with event data aggregates by months or even years, the HMM requires no temporal aggregation. This is important for modeling political behavior, where critical developments may occur over a week or even a day. The HMM is relatively insensitive to the delineation of the start of a sequence. It is simple to prefix an HMM with an initial "background" state that reflects the distribution of events generated by a source (e.g. Reuters/WEIS) when no crisis is occurring. A model can cycle in this state until something important happens and it moves into the later states characteristic of mediation behavior.

Data

The event data used in this study were machine-coded using the WEIS system from Reuters lead sentences obtained from the NEXIS data service for the period April 1979 through May 1997 and the Reuters Business Briefing for June 1997 through December 1998 using the Kansas Event Data System (KEDS) program (Gerner et al. 1994; Schrodt, Davis & Weddle 1994).[5] KEDS does some simple linguistic parsing of the news reports—for instance, it identifies the political actors, recognizes compound nouns and compound verb phrases, and determines the references of pronouns—and then employs a large set of verb patterns to determine the appropriate event code. Schrodt and Gerner (1994), Huxtable and Pevehouse (1996) and Bond et al. (1997) discuss

---

[5] The NEXIS "REUNA" file was used as the source for the period 15 April 79 to 10 June 97; Reuters Business Briefing (RBB) was used for the period 11 June 97 to 31 December 98. The change of sources was required because Reuters stopped supplying data to NEXIS on 10 June 97; the two data services provide a somewhat different mix of stories but there is no evidence of a discontinuity when the stories are coded and aggregated at a monthly level.

The following search command was used to locate stories in NEXIS :

(ISRAEL! OR PLO OR PALEST! OR LEBAN! OR JORDAN! OR SYRIA! OR EGYPT!) AND NOT (SOCCER! OR SPORT! OR OLYMPIC! OR TENNIS OR BASKETBALL)

To locate stories in RBB, the RBB search software (version 2.0 for Macintosh) was used to select the countries Egypt, Israel, Jordan, Lebanon and Syria; the "Israel" category includes stories dealing with the Palestine National Authority as well as Israel. The "Political" and "General" news sources were selected; the "Reuters Sports" source was excluded.

Some additional filtering was done on both the NEXIS and RBB downloads to eliminate Reuters "Highlights", historical calendars and other irrelevant material; details on this filtering are found in the archived data sets or from the authors. Only the lead sentences of the stories were coded; this produced a total of 92,687 events.

extensively the reliability and validity of event data generated using Reuters and KEDS. Figures 3 and 4 show the Goldstein-scaled time series for the USA>ISR and USA>PAL series.[6]

In order to deal with the dyadic character of the data (USA>ISR and USA>PAL), a 45-state model was used. The first 22 states are the 22 2-digit WEIS categories for USA>ISR. The USA>PAL events were assigned the codes **23** through **44**, corresponding to the original WEIS codes **01** to **22**. If no event occurred with either dyad during a day, a **00** nonevent was assigned to the day, so that each day has at least one coded event.

The terms of office of various U.S. Secretaries of States who during this period are given in Table 1. Four of these Secretaries have sufficiently long periods in office that it is reasonable to try to estimate a model: George Schultz, James Baker, Warren Christopher and Madeline Albright. While I will refer to each of mediation styles by the name of the Secretary of State during whose term that style occurred, this does not necessarily mean that the Secretary was involved directly in the mediation: for example Schultz tended to delegate much of his Middle East negotiating, whereas Christopher took a more hands-on approach.

Table 1.
U.S. Secretaries of State 1979-1998

Secretary	Start Date	End Date
Vance	23-Jan-77	28-Apr-80
Muskie	8-May-80	18-Jan-81
Haig	22-Jan-81	5-Jul-82
Schultz	16-Jul-82	20-Jan-89
Baker	25-Jan-89	23-Aug-92
Engleberger	8-Dec-92	19-Jan-93
Christopher	20-Jan-93	17-Jan-97
Albright	23-Jan-97	— —

Source: http://www.state.gov/www/about_state/history/officers/secstate.html

[6] Following the notational conventions of the KEDS project, "USA>ISR" refers to the series of actions by the United States towards Israel. "ISR-PAL" refers to actions by Israel towards Palestinians and actions by Palestinians towards Israel.

Figure 3. Goldstein Scores: USA>Israel

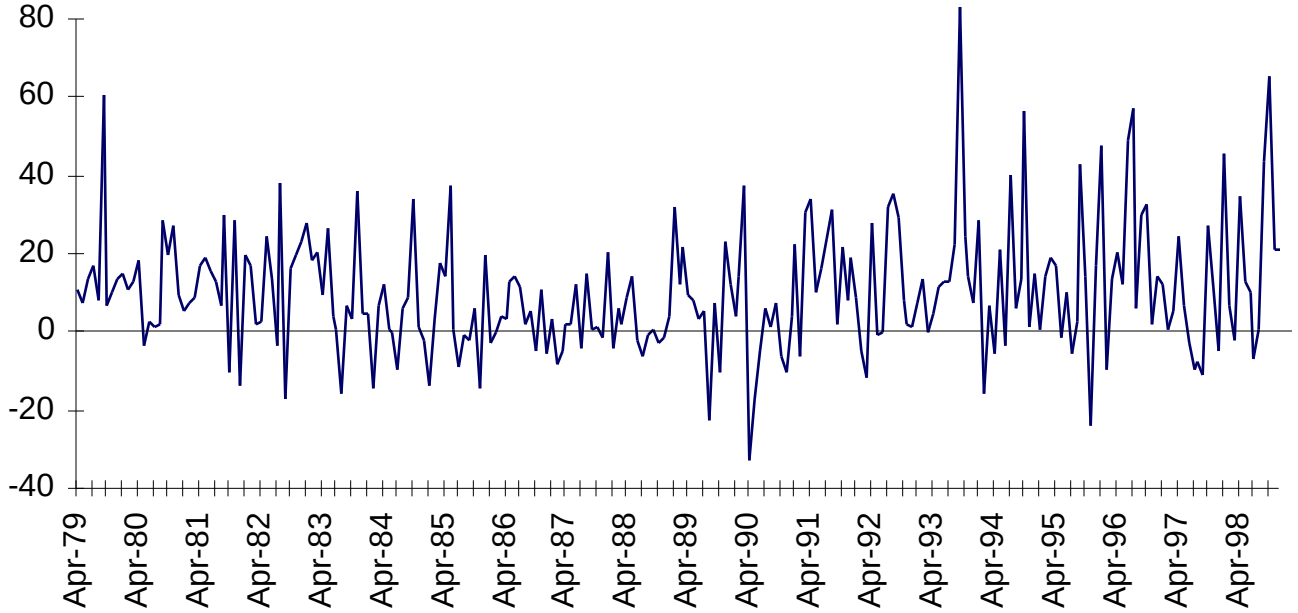
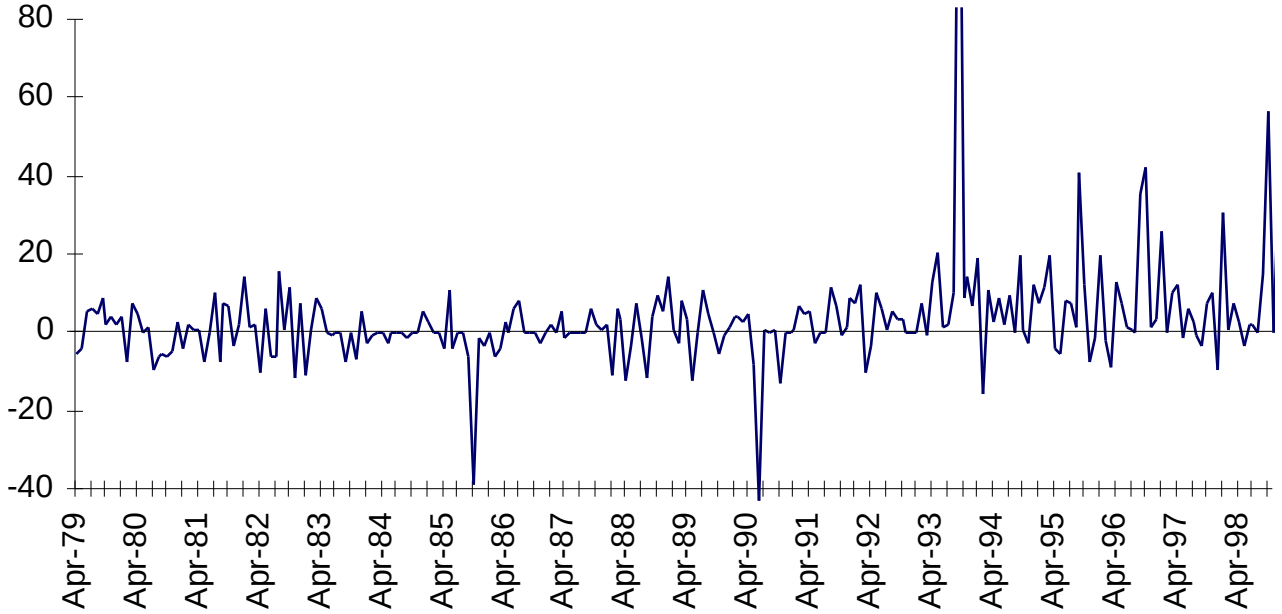


Figure 4. Goldstein Scores: USA>Palestinians



A clear potential problem with this analysis is that are changes Israeli-Palestinian relations over the 20-year time period contained in this data, and within the series, there are also distinctive periods such as the Palestinian *intifada*. Statistically speaking, the process is not stationary.[7] These changes may be due in part to changes U.S. mediation styles, but they could also occur independent of them: for example the single greatest shift in relations, the mutual recognition coming out of the Oslo Accords, occurred independently of any U.S. efforts. This does not preclude meaningful analysis—and it affects qualitative studies as much as quantitative—but it means that one cannot accept that all differences in behavior necessarily follow from changes in U.S. efforts.

Estimation Algorithm

The HMM parameters were estimated by extensively modifying the source code written by Meyers & Whitson (1995). Their C++ code implements an LR hidden Markov model and the corresponding Baum-Welch maximum likelihood training algorithm. I translated this code from the Solaris C++ environment to an ANSI C environment, in the process combining Meyers and Whitson's separate driver programs for training and testing into a single program, and modifying the input format to handle the WEIS sequences. The source code for this program is available at the KEDS web site: http://www.ukans.edu/~keds. I then extended the code to handle the LRL

[7] Somewhat to my surprise, there is little formal statistical trend in the ISR>PAL series when this is aggregated using the Goldstein scale: in fact technically the trend is negative. There is, however, a distinct trend in improving USA-Palestinian relations. The following are the results of regressing the Goldstein scores against time (monthly aggregation) for the dyads considered in this analysis

Dyad	Slope	Intercept	r	significance
USA>ISR	0.02	7.72	0.10	0.10
USA>PAL	0.05	-2.72	0.27	<0.001
ISR>USA	0.00	9.00	0.00	.99
PAL>USA	0.04	-2.92	0.21	<0.001
ISR>PAL	-0.21	-59.52	-0.17	0.006
PAL>ISR	-0.05	-24.86	-0.09	0.15

model, and implemented the Viterbi algorithm described in Rabiner (1989) in order to estimate the most likely state sequence.[8]

The resulting program is very fast—estimation of the HMM matrices using six 100-event sequences with a 45-symbol set and 64 Monte-Carlo iterations of the initial matrix takes about 45 seconds on a Power Macintosh 7100/80, and the computation of the probability of a sequence being generated by a particular HMM is nearly instantaneous. The program requires about 1 Mb of memory for a system using 45 codes, 6 states and 100-event sequences. The largest arrays required by the program are proportional to (M+T)*N, where M is the number of possible event codes, T is the maximum sequence length and N is the number of states. Consistent with earlier HMM work (Schrodt 1999), the models I estimated used 6 states; the six-state model has been widely employed in the studies cited earlier and appears to work quite well for differentiating crisis outcome and forecasting in the Middle East.

The models were estimated as follows. First, sixteen 32-event sequences were generated randomly from the term of office of each of the four Secretaries.[9] These sequences were created by generating a random date, then taking the 32 *events* prior to that date; note that this would involve a period of 32 *days* only if the sequence consisted solely of non-event codes, so usually the sequence will be shorter than 32 days. The best fitting model from 512 Monte-Carlo experiments—the model that matched the 16 training sequences with the highest total probability—was then used to characterize each Secretary. After these four models had been estimated, the Secretarial "style" that best characterized each month of data in the series was computed by taking the 32 events prior to the end of the month, and then determining which of

---

[8] The Meyers & Whitson code is clean, well-documented, and survived my translation to run correctly the first time. I would assume that either my C code or their C++ code would port easily to a DOS/Windows or OS/2 environment for those so inclined. In the process of extending the model to the LRL form, I rewrote the estimation equations to correspond exactly to those in Rabiner—the Meyers & Whitson implementation differed slightly from Rabiner's equations, presumably because their models estimate a separate vector for "transition symbols." These new procedures produce estimates similar to those of Meyers & Whitson when all probabilities to previous states are forced to zero. The program used here is a slight modification of that used in Schrodt (1999), which contains a much more extensive set of routines.

[9] The earliest starting date of a training sequence was 32 days after the beginning of the term of office.

the four HMMs had the highest probability of generating that sequence. Finally, this process was repeated 16 times and the modal (i.e. most commonly identified) style was assigned to each month in the series.[10]

Clear and unambiguous design, right? Well, after I had run the analysis, interpreted the results and presented the paper, I discovered the following statement in the code used to estimate the model

#define fLR TRUE // LR matrix?

This was a logical switch in the program that limited estimation to a "left-right" model; it had remained there from some experiments I had done for an earlier paper that compared LR and LRL estimates.

How common is this type of error in research? Very. In a now-classical study (Dewald, Thursby and Anderson 1986; see also King 1995) the editor of the economics-oriented Journal of Money, Credit and Banking *attempted to replicate the statistical results of a year's worth of articles in the journal, and failed completely in most cases. In highly computer-dependent research, mistakes like this are extremely easy to make.*

Because of this, most major social science journals now require authors to archive data sets so that scholars can duplicate the study in the future. Typically studies are not exactly replicated, but if a result is taken seriously, it will be studied using a number of different statistical methods, data sets and definitions of variables. Only if the statistical conclusions hold up in a variety of tests will it be accepted.

In this particular case, the mistake is probably harmless—the research I had done earlier showed that resulted obtained with LR and LRL models were quite similar. But generally one should not trust a result—whether in the social or the natural sciences—until it has been replicated.

Results

As shown in Figure 5, the HMMs differentiate the various mediation styles of the four Secretaries quite distinctly. The dots show the category that each month was assigned; the vertical dotted lines show the transitions between Secretaries. The assignment is not perfect, but

[10] Despite the large amount of computation, the whole process—16 repetitions of 512 Monte-Carlo experiments on 4 models—takes only about eight hours on a Macintosh 7100/80, and presumably would take substantially less time on a faster machine.

this was expected: at various points in time a mediator might be using a style that is similar to that of a different period.[11]

The "Baker" style is most distinctive, with only a small number of departures to the "Albright" style. The distinctiveness of this period might be due to the fact that Baker and Bush were closely involved in foreign policy (and had a relatively clear foreign policy agenda), though it could also be due to the effects of the Palestinian *intifada*. However, the Baker style ends quite abruptly, which may be due to the contrast between the often tense relations between the Bush administration and Israel, and the very warm relations that characterized the early Clinton administration. The other very distinct style is Albright's, though this might partly be due to the relatively short length of this period, which means that the sample sequences used to estimate the model covered much of the data.

In contrast, the "Schultz" style is the least consistent, containing a mixture of all four styles, albeit predominantly its own style during the middle of the period. The strong shift to the "Baker" style at the beginning and end of the term appears to be the consequence of Israel's invasion of Lebanon in June 1982 (thought this inter-mixes the Baker and Christopher style) and the outbreak of the Palestinian *intifada* in December 1987. Christopher generally mixes styles in a manner similar to Schultz.

The period at the very beginning of the series, corresponding to the Vance, Muskie and Haig terms in office, shows almost a uniform distribution of assignments (Schultz:9; Baker:7; Christopher :12; Albright:7). This suggests either a high level of inconsistency in policy during this period—a deduction consistent with the high turnover of Secretaries—or else that this period (or periods) had a separate distinct style and the estimated styles are being assigned at random.

---

[11]I also experimented with the same estimation design but using random sequences that were 100 events in length. As expected, this increases the accuracy of the assignment of style. However, since the tests of the effects of the styles involved looking at out-of-term behavior, I used the shorter sequences to increase the possibility of identifying short periods of time when a mediator was using a different style.

Figure 5. Styles in USA>ISR and USA>PAL data

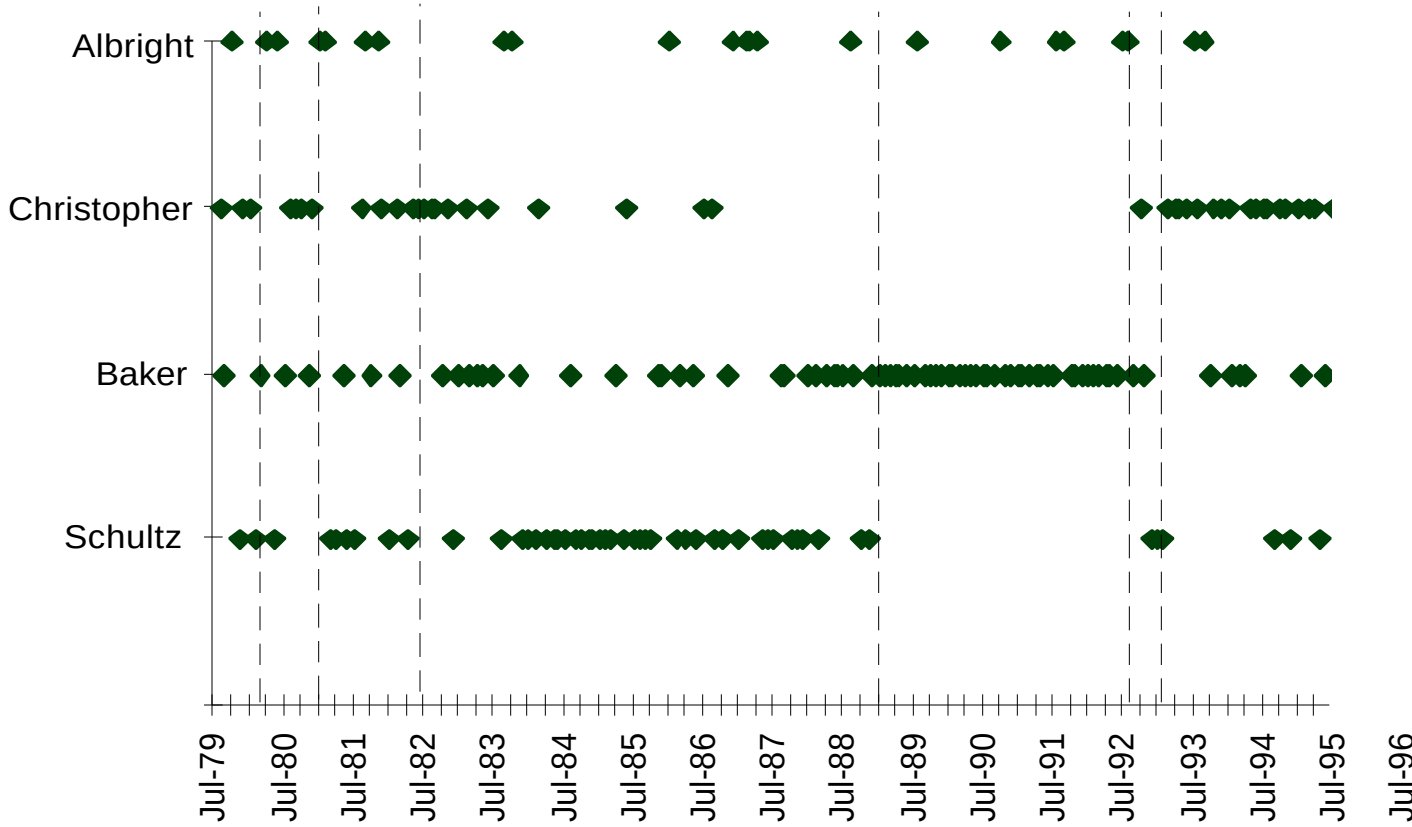


Figure 6. Frequency of modal style, USA>ISR-PAL

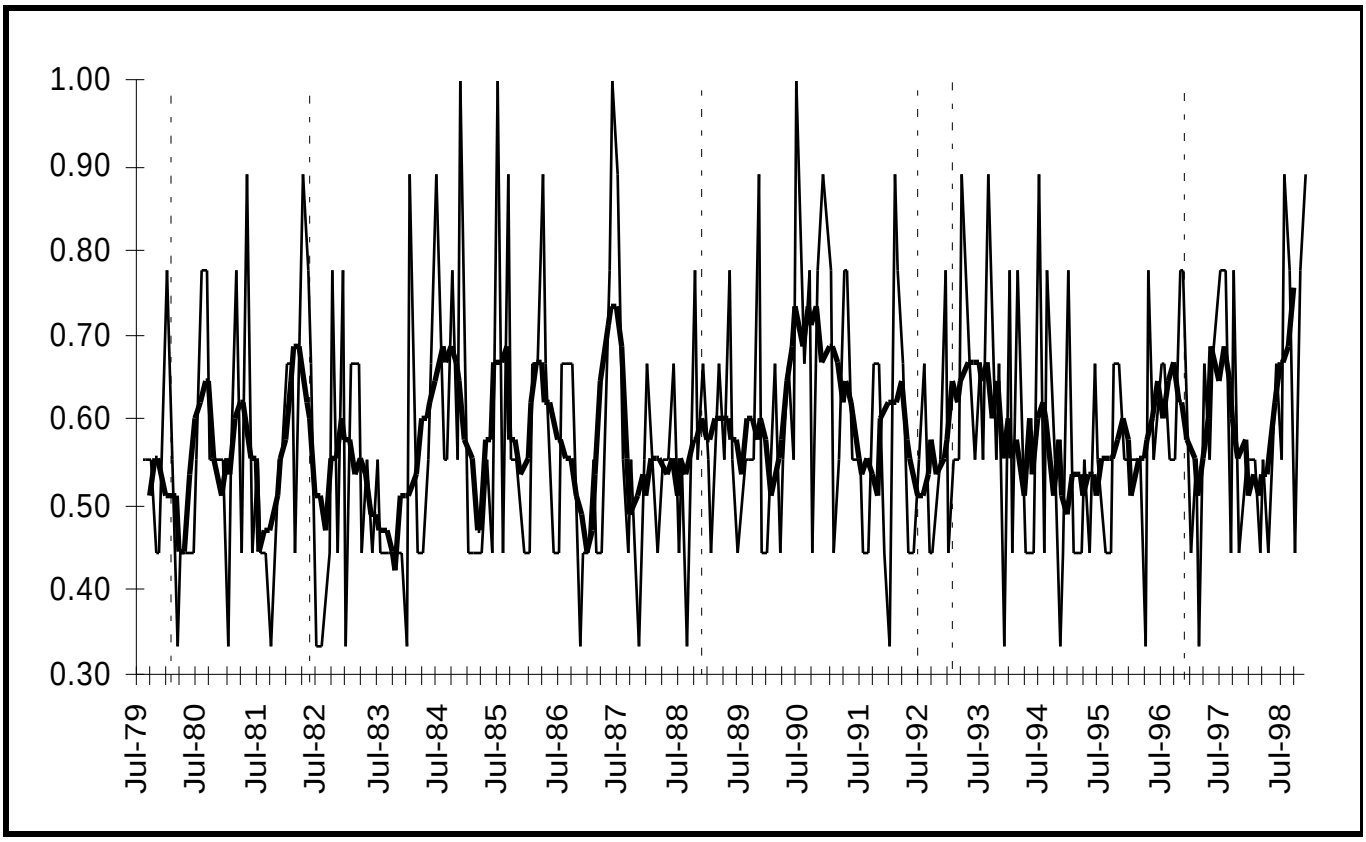
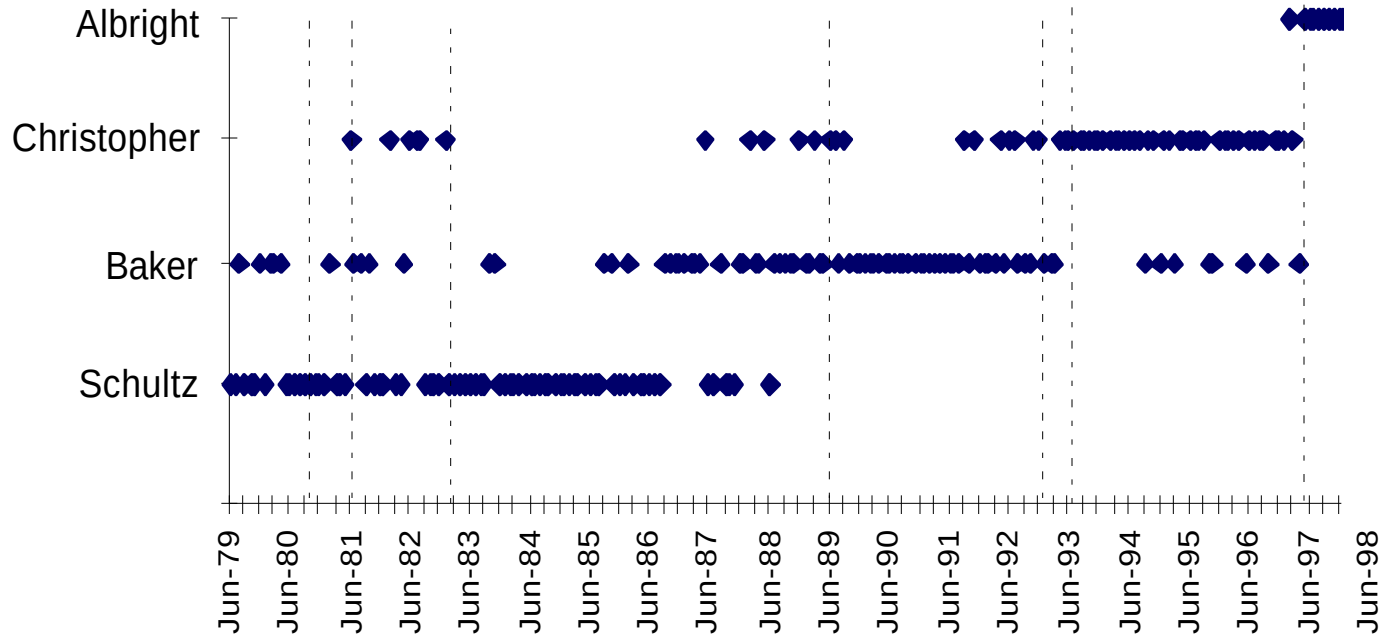


Figure 6 shows the consistency of the assignment of the modal category over time: the scores range from 0.25 (which would occur if the 4 styles were assigned with to a given month equal frequency in the 16 repetitions) to 1.0 (the month is always assigned the same style). The narrow line shows the monthly values of this measure; the dark line is a five-month centered moving average. The modal assignment is usually in 50% to 70% of the repetitions, a level well above what would be expected by chance, but otherwise few patterns are evident. The one possible pattern is the high level of consistency found in the middle of the Schultz and Baker terms, though the opposite pattern holds for Christopher, where there is a decline in the middle of the term. (The pattern for Albright is too short to generalize, though at the end of the data set the assignment confidence is reaching high levels.)

The fact that Israel-Palestinian relations are themselves changing over time opens the possibility that the ISR-PAL interactions are actually determining the styles. To test this possibility, I ran the same analysis on data from the ISR>PAL and PAL>ISR dyads, using the same periods as before. These results are shown in Figure 7.

These tests were time-consuming but absolutely critical: Because of the very large number of parameters in the HMM, it is possible that it would fit anything *with much the same pattern that I found with the Secretaries of State. The advantage of modern computers is that we can test very complex models; the disadvantage is the ease with which one can engage in machine-assisted self-deception.*

Figure 7. Styles in Israel-Palestinian dyad



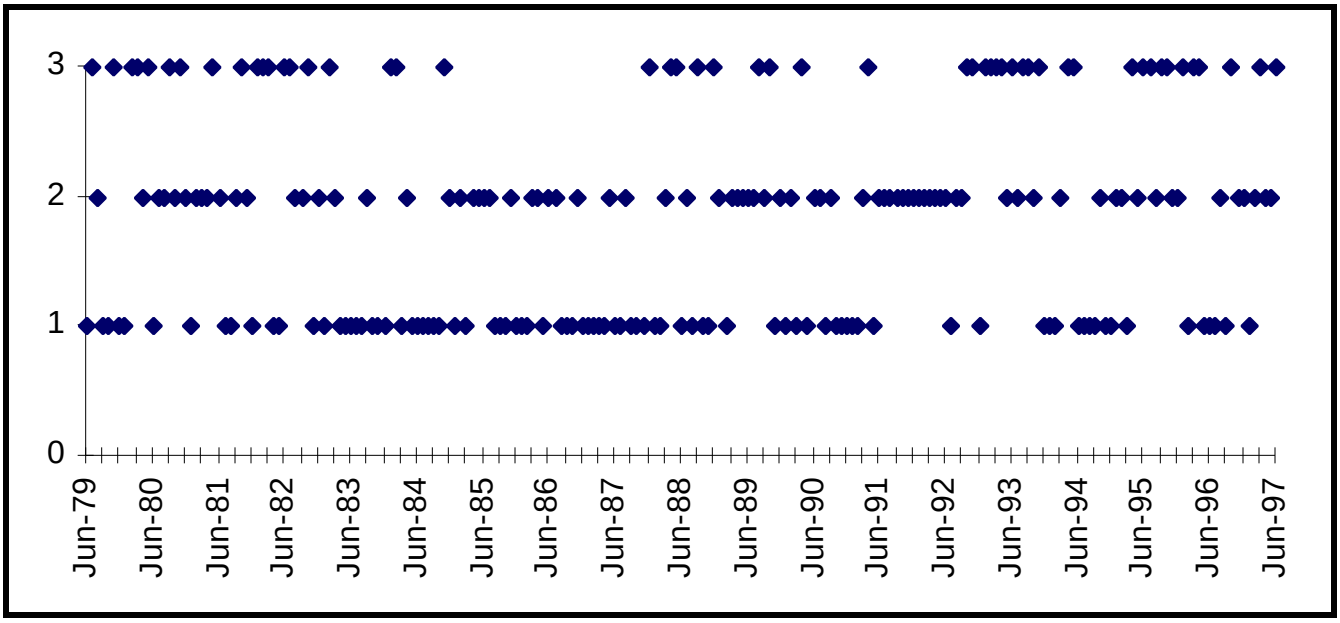
The patterns in Figure 7 is quite distinct from that in Figure 5 in two respects. First, Figure 7 shows much *greater* dependence of the style assignments over time than found in Figure 5: for example none of the months in the pre-Schultz or Schultz periods are assigned to the Albright category, and none of the months in the Albright category are assigned to Schultz. Second, there is a much less dependence on the shifts in Secretarial terms—for example the Baker/*intifada* period extends back through the beginning of 1987 (where the Goldstein-scaled event data begin to show a rise in violence prior to the *intifada*) and shows a very ambiguous, rather than an abruptly changing, classification at the end of the Baker terms. Similarly, the Schultz model is almost as strong in the pre-Schultz period as in later times. While the effects of Israel-Palestinian relations are having some effect on the assignment of clusters (as one would expect: this is after all the objective of the U.S. mediation), it does not fully determine them.

I'm going out on a bit of a limb here, and I've not done any formal tests. The absence of formal tests is due to two factors. First, this is an unconventional technique and there is no single obvious test. Second, most any test that did exist would test for the null hypothesis that there was no *relationship between the two sets of results, whereas I already know that there is at least* some *relationship. As outlined in the text, there are sufficient qualitative differences between the two figures that I'm convinced Figure 5 shows something more than just the effects of changing Israeli-Palestinian relations.*

Finally, in the interests of avoiding the possibility of computer-assisted ocular self-deception, the protocol was run on a set of data that were randomly ordered but had the same marginal frequency as the ISR-PAL data set. (This data set was generated for Schrodt (1997) and only goes through June 1997, so the Albright model is not included). These results, shown in Figure 8, are dramatically different than the patterns in Figure 5 and 7, and decisively rule out the possibility that this protocol will find clusters in any set of data.

Again, no numerical test here, but why bother? Sometimes the eyeball test is a reasonable approach. But, as noted below, it can also be wrong.

Figure 8. Styles in random data



The Effects of Mediation Style on Behavior

Having established that there are distinctive mediation styles, the next question is whether some of these styles are more effective at conflict reduction than others. In order to evaluate this proposition, I will look at whether there are changes in the average level of cooperation as measured by monthly aggregations of the Goldstein (1992) scaled values of the behavior within the remaining four parts of the USA-ISR-PAL triad.

Table 2 shows the t-test results (assuming unequal variances) for the average Goldstein scores in months characterized by various mediation styles (this includes the pre-Schultz and the Engleberger periods) for the ISR>PAL, ISR>USA, PAL>ISR and PAL>USA dyads. For example, the -1.02 figure in the Chr/Bak entry to the ISR>PAL table is the value for the t-test for the difference in the average Goldstein score for month classified as having the Baker style and months classified as having the Christopher style. All of the t-tests are done on the difference (earlier style) minus (later style)—for example Baker-style minus Christopher-style. (The

degrees of freedom for the t-tests in these tables are between 40 and 130, so the using the critical value of about 2.0 for a 5% significance level is appropriate[12])

> *I had originally expected that these t-tests would show very clear results, and my initial analysis was done with an elaborate Microsoft Excel spreadsheet. Spreadsheets are very useful for exploratory work, particularly in time series, where a graph can be created in a few seconds. However, the ease with which spreadsheets can be manipulated also hides a danger: it is very easy to make a mistake—most of mine occur when I rearrange cells in the spreadsheet and fail to modify some formulas—and the mistake is rarely obvious.*
>
> *I originally used the spreadsheet because I figured I would compute the t-tests only once and then go on. However, when it became clear that tests involving multiple leads, lags and data sets would be required, the spreadsheet became too awkward and I wrote a separate computer program in C to do the calculation instead. It took an hour or so to write the program, but once written, the parameters of the test were clear, and the program could run tests on six series in less than a second.*
>
> *[I checked these results against Excel and was surprised at the differences, which were typically in the second decimal place, a high level for this type of work. I'd heard earlier that this low level of accuracy was a danger when running statistical work in Excel but hadn't actually seen it before.]*

When all months are included in the analysis, the significant differences are found only in the reactions between the Middle Eastern protagonists and the USA, not between Israelis and Palestinians. This occurs with the Schultz-Albright, Baker-Christopher, and Baker-Albright style pairs for both the USA>ISR and USA>PAL.[13] However, none of the style-pairs are significantly different in the ISR>PAL relationship, and only the Baker-Albright pair is significant for PAL>ISR, a frequency of significant differences that barely rises above the level of chance.

When only out-of-term months are considered, the results are even weaker. When the contemporaneous differences are considered, only two are significant: Baker-Albright for ISR>USA and PAL>USA (this could be due to chance, though that is unlikely given that the same style-pair is identified in both instances). Explorations of the possibility that the impact of the mediation style might involve a time lag or time lead fared little better: Table 3 shows the t-tests on the differences for the average behavior in the month *following* the use of an out-of-term

---

[12] If you are looking for greater precision than this in evaluating the significance levels for these tests, you're in the wrong business...

[13] Okay, Schultz-Albright technically just misses, but this is very noisy data.

style, and only two of the 24 comparisons are statistically significant: Schultz-Albright and Baker-Albright for PAL>ISR These were the strongest results for leading or lagging differences; most of the tests were even weaker in only a couple of cases, probably spurious.

> *Here I reach a critical point: do I believe that the t-tests are showing real differences?. Technically, I could argue that the tests were exceeding the critical value of the t-test at a rate exceeding chance, even when accounting informally for the fact I was running many tests (there are actually formal ways to adjust for multiple significance tests, though I don't know anyone who uses them.) However, I hadn't seen any really strong patterns, and based on many years of "explaining" patterns to myself only to discover later that they were due to programming errors, I was very aware of my ability to create a "story" to explain any set of results.*[14] *How should I spin the interpretation?*
>
> *In the end, I decided to conclude that there was no pattern in the results. This was based largely on my previous experience with this data set, which I knew would often show a clear pattern of rise-and-fall in various statistical measures as one adjusts the lead and lags. The technical term for this is "autocorrelation"; in plain English, political organizations tend to do the same thing today and tomorrow that they did yesterday, so introducing delays in the relationship should show gradual, rather than sudden, changes in the results. This was particularly true in this design because a "month" was an arbitrary unit of aggregation.*

Some additional insight into this analysis can be found from Table 4, which repeats the t-test analysis on the cases dyads used to estimate the original models, USA>ISR and USA>PAL. Despite the fact that the these cases are differentiated quite nicely by the HMM technique, there are only two significant differences in each of the dyads when all of the months are analyzed, and *no* significant differences in the out-of-term months! This suggests that the differences in mediation styles are something more subtle than what can be measured through simple aggregated Goldstein scores.

---

[14] This problem is not unique to the social sciences. A number of years ago some radio astronomers thought they had discovered a totally new form of pulsar. The signals that led to this "discovery" subsequently turned out to be due to a malfunctioning computer chip and the result was withdrawn, but in the meantime a couple of dozen papers had been written to explain how this unusual pulsar could have been formed.

Table 2. All Months

	ISR>PAL				ISR>USA		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	0.44			Bak	0.52		
Chr	-0.49	-1.02		Chr	-1.81	-2.29	
Alb	-1.11	-1.68	-0.73	Alb	-1.90	-2.38	-0.43
	PAL>ISR				**PAL>USA**		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	1.90			Bak	0.80		
Chr	0.37	-1.39		Chr	-1.37	-2.10	
Alb	-0.91	-2.36	-1.10	Alb	-2.62	-3.45	-1.76

Table 3. Out-of-Term Months, 1 Month Lead

	ISR>PAL				ISR>USA		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	0.17			Bak	0.81		
Chr	-1.34	-1.64		Chr	-0.33	-1.33	
Alb	-1.42	-1.59	0.27	Alb	0.84	0.09	1.28
	PAL>ISR				**PAL>USA**		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	-0.48			Bak	0.45		
Chr	-1.52	-1.33		Chr	-0.39	-1.16	
Alb	-3.31	-2.73	-1.01	Alb	0.45	-1.29	-0.04

Table 4. T-tests on training sets

All Months

	USA>ISR				USA>PAL		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	-1.61			Bak	0.60		
Chr	-2.52	-1.22		Chr	-1.64	-2.49	
Alb	2.54	-1.45	-0.32	Alb	-1.62	-2.31	-0.62

Out-of-term Months

	USA>ISR				USA>PAL		
	Sch	Bak	Chr		Sch	Bak	Chr
Bak	-0.24			Bak	1.64		
Chr	-0.53	-0.38		Chr	0.71	-0.49	
Alb	-0.88	-0.96	-0.50	Alb	-0.45	-1.67	-0.99

Here we come to a point where I'm caught by two assumptions that I had made early in the project but not tested. The first was that the USA>ISR and USA>PAL series would show distinct differences when evaluated using t-tests on the aggregated data. Because the differences had been so clear in the HMM model—and because I did my original tests using a spreadsheet—I hadn't tested this. It's wrong.

Second, I had assumed—based on the now untrustworthy *"eyeball" test—that the Goldstein scaled data would indicate a significant trend in improving relations from 1979 to 1998. Again, wrong: When I finally ran the tests that were reported back in Footnote 7, only the USA-PAL relationship shows trend. Despite its location in the text, footnote 7 is one of the* last *statistical tests I ran on the data, and after finding the lack of trend, I went back through the text of the paper and removed a assortment of comments I had made that had explained results on the basis of trend.*

On the other hand, this failure opened an opportunity: the Goldstein scale had generally worked well, but now it failed, and the difference was probably due to the emphasis in the research on mediation rather than conflict. This becomes a major theme in the conclusion.

Conclusion

This paper is only a first small step in analyzing the characteristics and effects of mediation strategies. The hidden Markov models provide clear evidence that there are differences in the mediation styles of various U.S. Secretaries of State, a conclusion that is consistent with qualitative assessments of history. Furthermore, those styles can be detected in event data using

objective methods, a conclusion that was not obvious. The character and effects of those differing styles, on the other hand, are not obvious through the conventional event data analysis methods employed scaled and aggregated events.[15]

Part of the problem may be in the use of weighting schemes such as Azar and Sloan (1975), Vincent (1979) or Goldstein (1992). While these appear to work fairly well in many applications, there are a couple of clear problems with scaling. First, aggregating events is controversial: the "folk criticism"[16] of the Azar-Sloan scale is "3 riots equals a nuclear war." This debate goes back to the earliest event data discussions (e.g. Azar and Ben-Dak 1975; Azar, Brody and McClelland 1972) and has continued over time: see exchanges between Howell (1983) and McClelland (1983) or Vincent (1990) and Dixon (1990).

The uni-dimensional cooperation-to-conflict scaling might also be problematic—for example the USA-Canada or USA-Japan relationships are characterized by high levels of both cooperation and political conflict. To date this hasn't prevented the scaled data from being used successfully in a variety of studies, but that may be due in part to the fact that event data have been primarily employed to study highly conflictual situations such as the Cold War (Ashley 1980; Goldstein and Freeman 1990; Dixon 1986) and antagonistic behavior in the Middle East (Azar 1972; Azar et al. 1979; Schrodt and Gerner 1997, 1998) where "cooperation" is largely expressed as a reduction of conflict.

This study of differing mediation styles, in contrast, may provide an instance where scaled event data does not present an accurate picture of the underlying behavior. While the *political objective* of mediation is the reduction of conflict, that will not necessarily be the short-term result. For example, one obvious problem in evaluating the effects of mediation in the Israeli-Palestinian dispute is that successful conflict-reduction by the core actors in the dispute—the

---

[15] For at least not in simple difference-of-means tests. For elaborate models, such as the VAR analyses used by Goldstein and Freeman (1990) might still show differences.

[16] As in "folk theorem": I've heard this phrase many times over the years but I have no idea who originated it. The Azar-Sloan value for "inciting of riots" is equal to 44; "full-scale war" is 102.

Israeli government and the PLO/PNA—has frequently led to increased violence by extremists on both side, notably right-wing settlers in Israel and Islamic militants among the Palestinians.

This suggests that the next step in the analysis should be determining what characteristics of the triadic relationship is changing. One possibility would be to look at the differences in the HMMs themselves. Unfortunately, the 282 coefficients in the HMMs estimated by the Monte-Carlo method are highly indeterminant (see Schrodt 1997), so it is unlikely that much can be discerned from the HMMs themselves, even when a single set of training sequences is used. It may, however, be possible to use a more sophisticated method such as a genetic algorithm to generate more stable estimates. Whether this is possible depends on whether the indeterminance of the estimates is intrinsic to the structure of the model and data— for example an effect similar to co-linearity in linear models—or whether it is simply due to the inefficiency and irregularity of the Monte-Carlo estimation method. If it is possible to consistently estimate a small class of models characterizing each style, the differences in the mediation styles should be evident from the distribution of event categories in the symbol vectors.

This indeterminacy is clearly exacerbated by the low ratio of HMM parameters (272) to observed events (16 training sequences of 32 events = 512 events) in the test protocol. The relatively short training sequences were used to avoid the possibility of tautological classification: all months would be associated with the correct Secretary because all of those events had been used in estimating the model. However, having shown that it is possible to differentiate styles with relatively brief sequences, it might be useful to estimate more comprehensive models—possibly including the entire terms of office—to determine what types of events and event combinations actually characterize the various mediation styles.

Alternatively, the differences and effects of the styles might be evident in fairly simple disaggregations from the Goldstein scale. The obvious alternative is to break out conflict (WEIS **01** to **10**) and conflict (WEIS **11** to **22**) on separate scales in order to eliminate the problem of extremist acts of violence masking cooperation among the core actors. Another possibility would

be to introduce an element of time into the analysis by looking, for example, at the changes in the frequency of short sequences of events such as reciprocal cooperative events.

The assumption of the "Multiple Paths to Knowledge Project" was to give a groups of researchers who were specialists in applying a variety of disparate techniques a difficult problem to work on. Based on this analysis, they were correct in identifying mediation as a difficult problem to analyze. The hidden Markov analysis clearly indicates that distinct mediation styles exist in the approaches used by U.S. Secretaries of States, and that these can be detected in event data using objective methods. For now, however, the precise content and effect of those styles remains a puzzle whose solution requires methods more subtle than those that are conventionally employed in the analysis of event data.

The paper was presented at the International Studies Association meetings in Washington. The discussant, J. David Singer(Michigan), correctly noted that while I had provided evidence for the existence of distinct mediation styles, I had provided no information whatsoever as to what distinguished those styles. To which I could only agree—the failure of the aggregated measures meant that the results of the paper were fairly weak.

Wilkenfeld, Taber and I met briefly at the ISA meetings to plot future strategy. Curiously, our three projects were in somewhat similar situations—we had successfully implemented the initial phases of our designs, but the results were weaker than we'd anticipated. Further work was needed before we could start to compare the results of the different approaches, though we seemed to be on the track of something.

Bibliography

Allan, Pierre. 1980. "Diplomatic Time and Climate: A Formal Model." *Journal of Peace Science* 4:133-150.

Ashley, Richard K. 1980. *The Political Economy of War and Peace.* London: Frances Pinter.

Azar, Edward E. 1982. *Conflict and Peace Data Bank (COPDAB), 1948-1978.* Ann Arbor: Inter-University Consortium for Political and Social Research. (ICPSR Codebook no. 7767, second edition)

Azar, Edward E., and Thomas Sloan. 1975. *Dimensions of Interaction*. Pittsburgh: University Center for International Studies, University of Pittsburgh.

Azar, Edward E. 1972. "Conflict escalation and conflict reduction in international crisis: Suez, 1956." *Journal of Conflict Resolution* 16,2:183-202.

Azar, Edward, R.D. McLaurin, Thomas Havener, Craig Murphy, Thomas Sloan, and Charles H. Wagner. 1977. "A System for Forecasting Strategic Crises: Findings and Speculations About Conflict in the Middle East." *International Interactions* 3:193-222.

Azar, Edward E., Richard A. Brody, and Charles A. McClelland, eds. 1972. *International Events Interaction Analysis: Some Research Considerations.* Beverly Hills: Sage Publications.

Bartholomew, D. J. 1971. *Stochastic Models for Social Processes.* New York: Wiley.

Bloomfield, Lincoln P., and Allen Moulton. 1989. *CASCON III: Computer-Aided System for Analysis of Local Conflicts.* Cambridge: MIT Center for International Studies.

Bloomfield, Lincoln P., and Allen Moulton. 1997. *Managing International Conflict.* New York: St. Martin's Press.

Butterworth, Robert Lyle. 1976. *Managing Interstate Conflict,1945-74: Data with Synopses.* Pittsburgh: University of Pittsburgh University Center for International Studies.

Dewald, William G, Jerry G. Thursby and Richard G. Anderson. 1986. "Replication in Empirical Economics: *Journal of Money, Credit and Banking* Project." *American Economic Review* 76,4: 587-603.

Dixon, William J. 1986. "Reciprocity in United States-Soviet Relations: Multiple Symmetry or Issue Linkage." *American Journal of Political Science* 30:421-45.

Dixon, William J. 1990. "Interdependence and Cooperation in Foreign Policy Behavior: A Comment on Vincent." *International Interactions* 16:109-114.

Eisenberg, Laura Zittrain and Neil Caplan. 1998. *Negotiating Arab-Israeli Peace.* Bloomington, IN: Indiana University Press.

Gerner, Deborah J. 1991. *One Land, Two Peoples: The Conflict over Palestine.* Boulder, CO: Westview Press.

Gerner, Deborah J., Philip A. Schrodt, Ronald A. Francisco, and Judy L. Weddle. 1994. "The Machine Coding of Events from Regional and International Sources." *International Studies Quarterly* 38:91-119.

Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36: 369-385.

Goldstein, Joshua S., and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.

Guyatt, Nicholas. 1998. *The Absence of Peace.* New York: Zed Books.

Howell, Llewellyn D. 1983. "A Comparative Study of the WEIS and COPDAB Data Sets." *International Studies Quarterly* 27: 149-159.

Huxtable, Phillip A., and Jon C. Pevehouse. 1996. "Potential Validity Problems in Events Data Collection." *International Studies Notes* 21: 8-19.

King, Gary. 1995. "Replication, replication." *PS: Political Science and Politics.* 28,3: 444-451.

Kruskal, Joseph. B. 1983. "An Overview of Sequence Comparison." In *Time Warps, String Edits and Macromolecules,* ed. David Sankoff and Joseph B. Kruskal. New York: Addison-Wesley.

Lund, Michael S. 1996. *Preventing Violent Conflicts: A Strategy for Preventive Diplomacy.* Washington, D.C.: United States Institute for Peace.

McClelland, Charles A. 1976. *World Event/Interaction Survey Codebook.* (ICPSR 5211). Ann Arbor: Inter-University Consortium for Political and Social Research.

McClelland, Charles A. 1983. "Let the User Beware." *International Studies Quarterly* 27:169-177.

Myers, R., and J. Whitson. 1995. HIDDEN MARKOV MODEL for automatic speech recognition (C++ source code). http://www.itl.atr.co.jp/comp.speech/Section6/Recognition/myers.hmm.html

Quandt, William B. 1993. *Peace Process: American Diplomacy and the Arab-Israeli Conflict since 1967.* Berkeley: University of California Press.

Rabiner, Lawrence R. 1989. "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition." *Proceedings of the IEEE* 77:257-286

Sankoff, David, and Joseph B. Kruskal, eds. 1983. *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison.* New York: Addison-Wesley.

Schrodt, Philip A. 1984. "Artificial Intelligence and International Crisis: An Application of Pattern Recognition." Paper presented at the American Political Science Association, Washington, DC.

Schrodt, Philip A. 1986 "Set Prediction of International Events using a Holland Classifier." Paper presented at the International Studies Association, Anaheim.

Bennett, Scott and Philip A. Schrodt. 1987. "Linear Event Patterns in WEIS Data." Paper presented at the American Political Science Association, Chicago.

Schrodt, Philip A. 1989. "Short Term Prediction of International Events using a Holland Classifier." *Mathematical and Computer Modelling* 12:589-600.

Schrodt, Philip A. 1990. "Parallel Event Sequences in International Crises." *Political Behavior* 12:97-123.

Schrodt, Philip A. 1991. "Pattern Recognition of International Event Sequences: A Machine Learning Approach." In *Artificial Intelligence and International Politics*, edited by Valerie Hudson. Boulder: Westview.

Schrodt, Philip A. 1997. "Early Warning of Conflict in Southern Lebanon using Hidden Markov Models." Paper presented at the American Political Science Association, Washington.

Schrodt, Philip A. 1999. "Pattern Recognition of International Crises using Hidden Markov Models." in Diane Richards, ed. *Non-linear Models and Methods in Political Science* Ann Arbor: University of Michigan Press.

Schrodt, Philip A., and Deborah J. Gerner. 1994 . "Validity assessment of a machine-coded event data set for the Middle East, 1982-1992." *American Journal of Political Science* 38: 825-854.

Schrodt, Philip A., Shannon G. Davis, and Judy L. Weddle. 1994. "Political Science: KEDS—A Program for the Machine Coding of Event Data." *Social Science Computer Review* 12: 561-588.

Schrodt, Philip A., and Deborah J. Gerner. 1996. "Using Cluster Analysis to Derive Early Warning Indicators for Political Change in the Middle East, 1979-1996." Paper presented at the American Political Science Association, San Francisco.

Schrodt, Philip A., and Deborah J. Gerner. 1997. "Empirical Indicators of Crisis Phase in the Middle East, 1982-1995." *Journal of Conflict Resolution* 41:529-552.

Schrodt, Philip A., and Deborah J. Gerner. 1998. "Cluster Analysis as an Early Warning Technique in the Middle East." pp. 95-107 in John L. Davies and Ted R. Gurr, eds. *Preventive Measures: Building Risk Assessment and Crisis Early Warning.* Lanham MD: Rowman and Littlefield.

Sherman, Frank L., and Laura Neack. 1993. "Imagining the Possibilities: The Prospects of Isolating the Genome of International Conflict from the SHERFACS Dataset." In *International Event-Data Developments: DDIR Phase II*, ed. Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes. Ann Arbor: University of Michigan Press.

Tessler, Marc. 1994. *A History of the Israeli-Palestinian Conflict.* Bloomington, IN: Indiana University Press.

Vertzberger, Yaacov I. 1990. *The World in their Minds: Information Processing, Cognition and Perception in Foreign Policy Decision Making.* Stanford: Stanford University Press.

Vincent, Jack E. 1979. *Project Theory.* Lanham, MD: University Press of America.

Vincent, Jack E. 1983. "WEIS vs. COPDAB: Correspondence Problems." *International Studies Quarterly* 27:160-169.

Vincent, Jack E. 1990. "Interdependence Reexamined." *International Interactions* 16:91-107.