November 2006

Enhancing decision analysis models with web agents

Jeffrey M. Keisler
University of Massachusetts Boston, jeff.keisler@umb.edu

Wei Zhang
University of Massachusetts Boston

Enhancing decision analysis models with web-agents

Jeffrey M. Keisler and Wei Zhang
Management Sciences and
Information Systems Department
College of Management
University of Massachusetts Boston
100 Morrissey Boulevard, Boston, MA 02125
jeff.keisler@umb.edu
wei.zhang@umb.edu

Abstract

This paper explores the idea of enhancing decision analysis by taking advantage of vast, real-time data available from the World Wide Web (the Web). We illustrate the idea by linking influencing diagrams with electronic agents that can utilize the Web in a very active way. Essentially, at the time of modeling, the result of an agent's actions is treated as a stochastic event. Probability distributions for nodes in the influence diagram are assessed conditioned on the range of outcomes for these events. When the influence diagram is evaluated the agent performs actions as defined by the model and by the state of nature. Structuring the links presents technical challenges including programming and decision analytic assessment. Such agents can interact with the Internet operating in ways analogous to probes, sensors, monitors, beacons or in other roles. If information is costly, mechanisms for information acquisition decisions are needed. The paper discusses managerial decision classes that are especially well-suited to this type of application. In these cases, it may be effective to use human-intensive approaches from decision analysis consulting practice to structure models while creatively using autonomous agents to generate experimental data. The concept is illustrated with small, non-technical examples.

Keywords: Decision analysis, agent, data mining, knowledge management, influence diagram.

1. Introduction

Decision analysis may be viewed as an approach to avoiding waste in the use and processing of information and in this regard is not well tuned to the Internet era. The challenge of decision-making has shifted from dealing with information scarcity to dealing with information abundance. Even though the Internet and the World Wide Web now provide a vast amount of data, it's difficult for decision makers to systematically utilize that data, largely because most web data are stored in unstructured textual format. Researchers have envisioned that technologies such as the semantic web (Berners-Lee, Hendler and Lassila 2001) would bring better structure to Web data and thus make the data more meaningful. However, the semantic web is not coming along as fast as many have hoped. Effectively utilizing web data in decision making remains a challenge.

With structured data, e.g., data stored in relational databases, searches yield information in a straightforward fashion. Using the more amorphous information now available through the Internet presents a sort of information retrieval problem: how to find and get the relevant information from a large corpus of text (Kowalski 1997). Fortunately, a large portion (though by no means all) of this information is incorporated and continually updated (by spider software) in storage locations operated and made accessible by the providers of the major search engines, e.g., Google or Alta Vista. Employing services offered by the search engines (Mueller 2004) becomes a viable solution to handling the Web data. In this sense, a synthesis of two approaches – web searching and decision analysis – suggests an approach to inform decision models that will not be overwhelmed by data.

In addition to their sheer volume, the Web data are dynamic and changing in nearly real-time since they are managed and offered by numerous providers. New data can quickly make their way to the Web. Effective handling of the real-time Web data can enhance the timeliness of decision-making. Decision support based on real-time data is not a new idea. Hess et al. (2000) describe components of a system that uses distributed databases and even distributed models for manipulating the data. Uses for real time data in decision support include competitive intelligence (Chen, Chau and Zeng 2002), customer-relationship management based on historical data (Bowman and Narayandas 2004), and customer recommendation systems (Ariely, Lynch and Aparicio 2004). In other contexts, real-time information allows dynamic control of strategies, as in program trading (Furbush 1989). The use of real-time data within pre-defined structures seems most appropriate for repeated and persistent decisions. Current versions of spreadsheet packages such as Microsoft Excel support live-feed data, so decision model development in this direction faces a minimal technical hurdle. Nevertheless, the use of real-time Web data, which are mostly unstructured, has been less developed. In contrast, traditional decision analysis as a mechanism for decision support often requires extensive human interaction in order to capture expert knowledge in a useful way. Thus,

it has developed useful, usable techniques for mapping knowledge. Furthermore, the focus of decision analysis on obtaining decision-relevant information is all the more relevant in an era of speeded up and automated decision-making.

We propose that software agents that search the Web can help to incorporate unstructured, constantly changing Web data into DA models. Software agents are programs which perform tasks on behalf of people including sending and receiving information (Bradshaw 1997). A computer-based decision model can on its own launch agents which will perform on the web some tasks that people had to perform. We shall discuss several types of agents, how they could operate, how they could support DA modeling, and how to construct them.

In summary, we intend to enhance traditional decision analysis models by effectively utilizing the Web data. This approach extends traditional decision analysis in two ways. First, it allows us to incorporate real-time Web data in a decision analytic framework and deal with the complexities of doing so, i.e., formalizing the approach for linking the Web data with subjective assessments. This opens the door to including as much real-time data as desired in as complete a decision making context as desired. This leads to the second extension, utilizing the Web data by querying/searching the Web. We view a query to the Internet as an event with an uncertain outcome. Data are obtained through a continuum of methods, from gathering prepared reports to seeking search results to active measurement to generating new experimental data. Thus, instead of using Monte Carlo simulation to represent our uncertainty, we may be able to just have nature (the Internet) run the experiment – a decision analyst's dream. This capability is growing rapidly, with the advent of such devices as polling sites, survey sites, myriad search engines, and intelligent agents.

For example, if we are concerned with passage a new law such as the repeal of the estate tax in the United States, a typical model might require an expert assessment of the probability of passage, p . If the model is to be used in the future, it will need an updated probability estimate and the expert must be consulted again. An alternate approach is to run the real-time specific experiment of conducting a web search for the phrase "estate tax repeal" in pages in the last month using Alta Vista and recording the number of hits. In constructing the model, the decision analyst would ask the expert to estimate $p|A$, the probability of repeal in the event that the number of hits is above 1,000 and $p|B$, the probability in the event that it is below 1,000. Then whenever the model is evaluated, an agent conducts a search and reports back its results, and the model uses either $p|A$ or $p|B$ as its probability depending on the search result.

Whether or not this is actually an improvement will depend on the extent to which there are experiments whose results really do correlate with the values in question and whether the assessment process can identify those correlations. In a future where there are 1,000 stories about the "pending defeat of the estate tax repeal," the model using the web

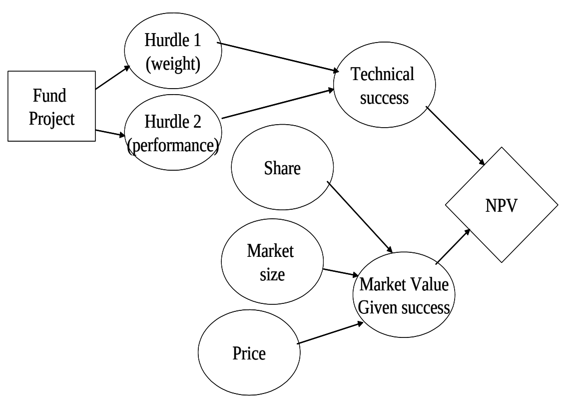
search would only note the last three words, and would give misleading results. The idea may be straightforward, but the methodology is not.

In this paper, we explore and illustrate how this concept could be used in web-based decision analytic support systems. For the purpose of illustration, we use influence diagrams as the front-end to these systems, allowing users to define the "experiments" with a guarantee of relevance and usability; to calculate optimal decisions, influence diagrams are often converted into equivalent decision trees. Software agents then connect models to the external world, making the models part of an open system. The modeling process requires interactions between agent design, decision analytic modeling and knowledge engineering.

We shall consider a decision problem typical of what business decision analysts see, and construct proof-of-concept tools that supplement the traditional model. Of particular interest are queries the results of which can be used to condition probability assessments. This effort raises issues and new possibilities we shall also discuss.

We start with a motivating example. Figure 1 shows the influence diagram for a hypothetical decision that would be a typical subject of decision analysis done by a consultant for a client (as in Howard 1988 and Howard 1989).

Figure 1: Influence diagram for a canonical example.

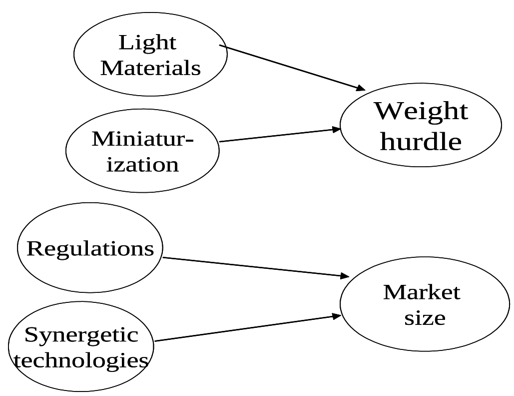


In this case, the company must make a decision about whether to fund a research and development (R&D) project, where the ENPV is calculated by computing the probability of technical success for the new product and the market value given success. The market value is derived from some standard accounting structures, and the probability of technical success is derived from the probability of overcoming a set of technical

hurdles. For example, when the technical hurdles are probabilistically independent, the probability of technical success is simply the product of the probabilities of overcoming all the individual hurdles. Here, we imagine that an automotive manufacturer is considering investing in development of a fuel-cell based engine. Real fuel cells are still heavy and, in practice, have some performance deficits (e.g., acceleration) Technical success here consists of achieving adequate performance with a light enough engine.

To construct a model like this one, the decision analyst would assess the chance nodes in conventional fashion moving leftward, and would typically halt when distributions can be assessed directly. For the leftmost assessments, it is common to use evocative nodes that remain unassessed, but that help structure the decision maker's thinking about the probabilities that are assessed. For example, in order to overcome the weight hurdle (by cutting weight 75% from the current state of the art) it would be necessary to develop materials that are 50% lighter per unit volume than the current best available and use these to construct an engine that is 50% of the volume than current technology allows. Similarly, the market size (which influences market value for the project) will depend on, among other things, whether regulations will require this technology for some products, and whether the technology is leveraged by other developments (e.g., availability of gas stations for fuel cell cars).

Figure 2: Unassessed evocative nodes implicitly influencing assessed nodes.



When such evocative nodes are used, experts may discuss them qualitatively before incorporating them into the probability assessments. For example, an expert on materials may have seen several articles describing promising new polymers, an expert on engine design may be aware of reductions achieved by a competitor, an expert on regulations may be aware of the history of a series of bills that are in some way related to the matter at hand. An expert on the overall market might be aware of development of, say, an operating system for electronic vehicles.

With the advent of the Internet, individuals do not just rely on rumors and personal networks. It is often practical and convenient for experts to quickly check the Internet in preparation for assessments. For the examples above, this could include: searching for a listing of articles, searching patents (Widing *et al.*, 1994), checking a trade journal forecast, or checking a competitor's news announcements.

In some situations, automated agents might do what these human experts do. The obvious benefits of using agents – high speed and low cost – can improve the ability of decision analysis models to be responsive and relevant for some applications. There are certainly precedents for this idea. Tseng and Gmytrasiewicz (2002), for example, constructed a decision support system for a well-defined domain (investment management) where agents query a well-defined set of source data in real-time. We can extend this idea, taking advantage of influence diagrams' inherent flexibility.

The construction of decision structures such as influence diagrams or value hierarchies (Keeney 1992) generally proceeds by constructing a quantitatively defined chain linking fundamental objects of concern back to directly observable conditions in the real-world. In this paper we shall consider only influence diagrams, but applying the same concepts to value hierarchies[1] should also be beneficial. Influence diagrams (Howard 1989) are constructed by asking "in order to predict node X, what one question would you most like answered" and constructing a predecessor node Y representing the uncertain answer to that question. The process is complete when the decision maker feels that the probability distribution for the last node can be adequately assessed directly, that is, there are no further uncertainties on which the assessment needs to be conditioned that would change the distribution due to finer judgments, or which might represent individually obtainable chunks of information whose value needs to be considered.

The tradeoff is that excessive detail requires exponential time, and possibly exponential judgmental complexity (too many conditioning variables to hold in one's mind at once), and therefore tends to introduce error. Ravinder *et al.* (1988) identified conditions where decomposing a problem can actually reduce accuracy. Howard (1989) leaves the door open to models run separately from the influence diagram and used as inputs to the influence diagram, in order to avoid problems such as cycles in the graphical network. This is not uncommon in practice, e.g., utility network utilizations are anticipated using a linear programming model, then summary statistics feed into a more standard decision analytic model. A common guideline in constructing decision models that combine

---

[1] Value hierarchies start with fundamental objectives and then chain back (asking how an objective could be measured), or start with more observable measures and move forward to identify their relation to fundamental objectives. Actual measures can vary greatly, e.g., geographic information system output has been used as model input. Elsewhere, constructed measures can quantify phenomena like "smell" with a scale using text (1 = like a skunk, 10 = like a rose, etc.) to show what each possible score means). The chains from measures to value eventually, of course, must be structured formally.

judgment and calculations is to judge that which is easily judged and compute that which is easily computed (and, presumably, to pray for the wisdom to tell the difference). In our case, decision analysts will define measures knowing that agents are available to measure.

2. Conceptual development: What can software agents do?

In order to use such generally defined material that agents may find, we must structure connection points between the agent and the influence diagram. That is, our standard influence diagram has nodes for which conditional probabilities are assessed for events where we would be informed by the "outcome" of the agent's activities. It is not unusual to have experiments in decision analysis and to condition probabilities based on their outcome, e.g., testing quality of a commodity product. By interpreting the actions of agents to be such experiments, we are able to be much more creative in designing the experiments – i.e., we can ask our agents to do anything! Much of the remainder of this paper will discuss the types of things we could ask them to do, how they would do them, and how a decision analytic model would incorporate the results.

Returning to our example, let us consider global warming regulations, noting that a probability distribution over size of the market for fuel cells could be assessed conditionally based upon the future state of Corporate Average Fuel Economy (CAFÉ) standards. Using Howard's clairvoyant test, we would operationalize this term, for example, as fleet average miles per gallon in the year 2010. In trying to predict CAFÉ standards, experts might ask a number of qualitative questions, and these could appear as evocative nodes – although in practice there might be too many to put them all down on paper. For example, the predicted level of future standards would be influenced by information about: the future political situation in the U.S. (will Republicans or Democrats control various parts of the government); the outcomes for related regulations (air pollution); and the development of scientific consensus regarding the severity and the causes of global warming.

For each of these evocative nodes, it is not difficult to identify dynamic electronic data that could be located by Internet agent at future points in time. An agent could go to any number of government sites and track the number of senators and representatives of each party (as well as the party of the president). Or, looking farther forward, the agent could retrieve data such as the current share prices in a predictive market (Berg and Reitz 2003) such as the Iowa Election Markets (IEM, http://www.biz.uiowa.edu/iem/) that indicate each party's probability of winning the next presidential election. In considering the outcome for related regulations, it would be possible to search a legal database to see which of a pre-defined set of bills (encoded by their standard identifying numbers at the time the influence diagram is constructed) has passed and when.

Regarding climate change, some models predict that frequency of hurricanes is an early warning sign of warming, and a future agent could go to a hurricane registry and tally the number (and strength) of hurricanes for the twelve month period immediately preceding. These are just examples, of course. It would be possible to assess the probability of a given level of CAFÉ requirements on any set of numbers that the agents retrieve. In practice, humans have difficulty performing conditional assessments when there are more than a small number of directly conditioning nodes. Therefore, we would want to define agents' functions well. We next discuss four different types of agents based on the specific types of data that agents might obtain to support decision making.

The most basic type of agent monitors sites that contain data in a persistent structure – the presidential vote share price for "Bush" at IEM, the average monthly temperature for Boston at weather.com, etc. (as in Tseng and Gmytrasiewicz, 2002).

Agents of this type, *monitors*, are relatively passive, and are fine for accessing such structured data as prices from markets, collated statistical reports, etc. Implementation of such agents is straightforward and their integration with decision models is typically straightforward. Current popular spreadsheet programs, for example can contain live data feeds updated by a user macro, and the results are automatically incorporated into any model that refers to the data location within the spreadsheet. Because the actual numbers an agent will return are unknown at the time the model is constructed, the agent's action is a type of experiment yielding an uncertain outcome.

At a slightly higher level of sophistication are agents that return a statistic that is algebraically or otherwise algorithmically derived from such experimental outcomes, e.g., quantities such as trends, ratios or correlations among the data are described by Bhargava and Power (2001). Limited only by our own creativity, we can derive from awkward raw material a set of statistics that are interesting and intuitive.

Within a given organization, it would be possible to leverage monitoring capabilities by creating *beacons* – web sites or even information portals containing potentially useful data that are made visible to search engines so that they can be found and used as needed by whatever models are constructed later.

While monitors and beacons involve structured data, more active agents – sensors and probes – cope with unstructured, mostly textual Web data. They can, for example, generate webometrics, i.e., quantitative summary data about web pages (Bjornebom and Ingwersen 2004). S*ensors* are agents that generate new data through non-invasive measurements. For example, the data obtained in the course of running of a web-query could include the number of hits for the query rather than simply the content of one of those hits. This doesn't change any visible data outside of the model. Obtaining such secondary data would likely require agents to use programmed queries including some special commands within a query language or even a more general programming language such as Java. Here, the experiment is telling us about the universe of

knowledge, rather than about a particular point. As with monitors, sensors can also return more sophisticated statistics derived from basic retrieved data.

More pro-active (Woolridge & Jennings, 1995) still would be *probes*. These agents are potentially invasive, that is, they do not leave the external world exactly as they found it – and thus may not have entirely replicable results. Probes provoke human or automated actors in the environment and track their response. For example, a probe could send a prepared email to a list and count the number of replies. More intelligent probes could have more complex interactions with the environment. Because they affect the environment, probes may bias future results from any type of agent, e.g., numerous probes about the airline ticket price for a specific flight may lead a vendor (or its automated program) to conclude that demand is increased and therefore to raise the ticket price. Thus, when incorporating probes, it's important not to ignore the possibility of such effects.

3. Implementation: Connecting agents to influence diagrams

We now develop some of these agents in a set of illustrative proof-of-concept examples related to the decision problem discussed earlier. The actions of these agents were performed by hand, rather than fully automated, with the purpose of finding out whether useful searches can be defined. The electronic implementation of the examples here would not be difficult, and specific languages and other details will depend on the specific applications developed.

Achieving a high level of improvement in fuel cell weight will require some substantial improvements in miniaturization. A basic sensor (i.e., one that does not derive any further statistics beyond what the search returns) ran a search in the AltaVista search engine for the sites, updated within a given time period, containing both the terms "Fuel Cells" and "miniaturization." Thus, the decision analyst in assessing the probability for the "Weight Hurdle" node would determine the expert's subjective probability that weight reduction (WR) would exceed 10%, 25% and 50% conditional on a search returning 10 hits, 100 hits, 1000 hits, etc., with a full distribution interpolated using common methods. For example:

$\Pr(WR > 10\% \mid Hits < 10) = 5\%$;

$\Pr(WR > 10\% \mid 10 < Hits < 100) = 15\%$;

$\Pr(WR > 10\% \mid Hits > 1000) = 30\%$, etc.

In fact, a search specifying sites constructed in 2003 returned 258 hits (note, this number itself will change over time as sites are catalogued), so in this case, that would mean that there is an approximately 20% chance that weight reduction would exceed 10%.

It is unlikely, however, that any expert would have a good intuition for how many hits indicates a high level of activity. Here, by having the agent perform two searches and deriving its reported statistic from them, we can alleviate that problem. Specifically, the agent looks for the increase in the number of hits, rather than the absolute number of hits. It finds that for the corresponding period in 2002, there are only 113 hits, and reports that there was an increase of 128% in activity in the last year. Again, the assessment for the weight hurdle node would be conditioned on statistic returned, e.g.,

pr(WR > 10%|Incr in hits < 20% = 5%);

pr(WR > 10%|Incr in hits >100% = 40%).

This is a much more flexible approach, but we have to be careful. How do we know in this case whether that increase is meaningful? Our expert may be fooled by the absolute number. To reduce the risk of this happening, the agent can create a control search looking at a similar statistic, and then compare the statistic for the control against the statistic about the phenomenon of interest, ultimately reporting a result that is derived from both. Using as a control a search for sites containing three randomly selected words "blue" and "liquid" and "chair" we find that the number of hits grew by 142% in the same time period – which renders the 128% figure less impressive. The statistic reported would be the ratio of these two numbers. It would take only slightly more work to implement this agent than the previous one. A decision model that could make use of this new finding would require assessments conditioned on the new statistic, e.g.,

pr(WR > 10% | {ratio of incr hits for "fuel cells" and "miniaturization" to increase in hits for secular search } < 1) = 2%.

In order to facilitate probabilistic assessments, experts may benefit from viewing empirical data regarding a variety of current and historical statistics for the phenomenon in question and similar phenomena. Hence, the process of construction is iterative, with the analyst/programmer feeding such results back to the expert before the expert assigns final numbers.

This example used only a single conditioning (evocative) node. To add information about lightweight materials, the agents would be similar but the assessments would be more complex, as the probabilities would be conditioned on two variables, although this is still well within standard decision analytic practice. When we can formulate a functional form of the relationship between the variables, e.g., weight = volume x density, we may substitute calculation for complex assessments.

Sensors can use more sophisticated search structure. We now shift to the global warming node of our influence diagram (which influences the market size node). We are not interested only in the level of discussion about global warming, but also whether the evidence indicates that global warming is for real. Here, the agent searched for the number of sites (6,140) containing the terms "global warming" and "conclusive" and

compared it to the number of sites (3,610) containing the terms "global warming and inconclusive." It is not difficult to generate possible search terms and it would not be taxing to test them to some extent at the time at which the influence diagram is constructed. Presumably, a search which used substitute terms such as "strong evidence" in place of "conclusive" and "weak evidence" in place of "inconclusive" would have generated similar, though not identical results. The percentage of hits on the side of conclusive evidence in our example is 6,140/9,750, but this is not, of course, our estimate of the probability that global warming is for real.

We implemented a prototype system using off-the-shelf decision analysis spreadsheet add-ins, along with VBA code to conduct web searches and extract metadata of the search results. In the cells containing probabilities and payoffs for the nodes of the decision tree (which can be automatically generated from influencing diagram), we changed the entries from constants to formulas which referenced cells containing the processed metadata. Figure 3 presents the screenshot of the decision tree. The probability in Cell J14 is tied with the search results presented in another worksheet (Web Search Tool in Figure 4). In this particular example, the probability is estimated by examining the average annual increase in hits resulted from searching the web.

Figure 3: Determining the probabilities in decision tree

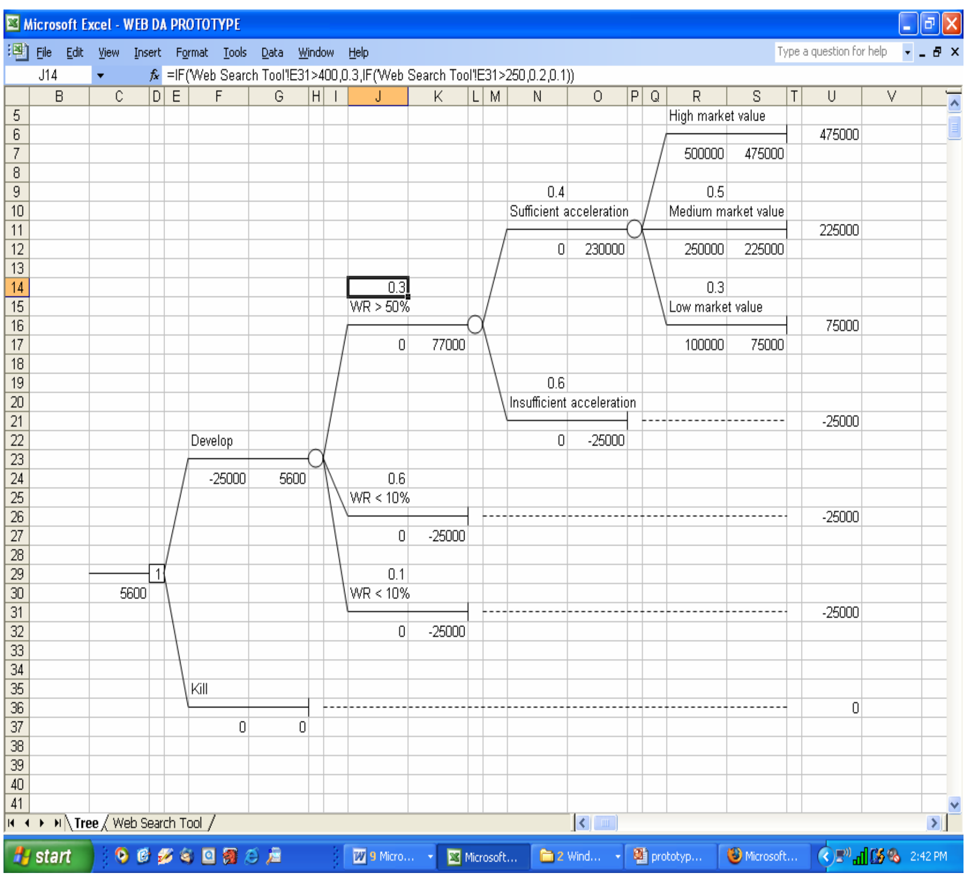
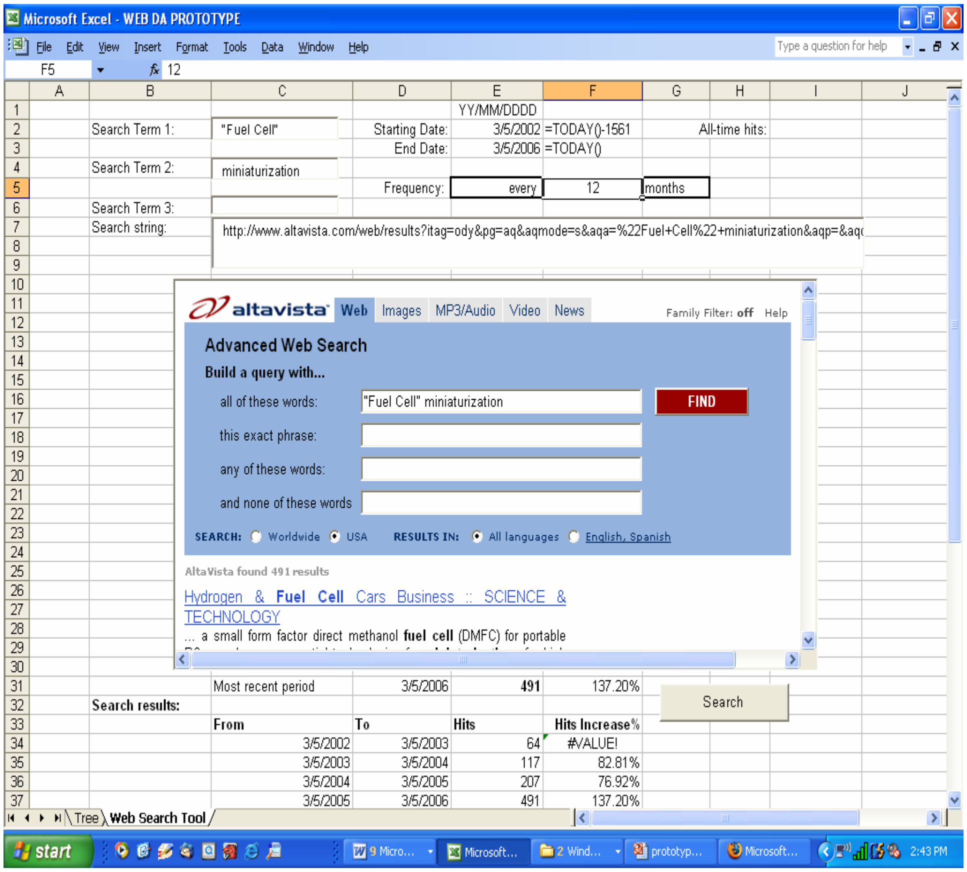
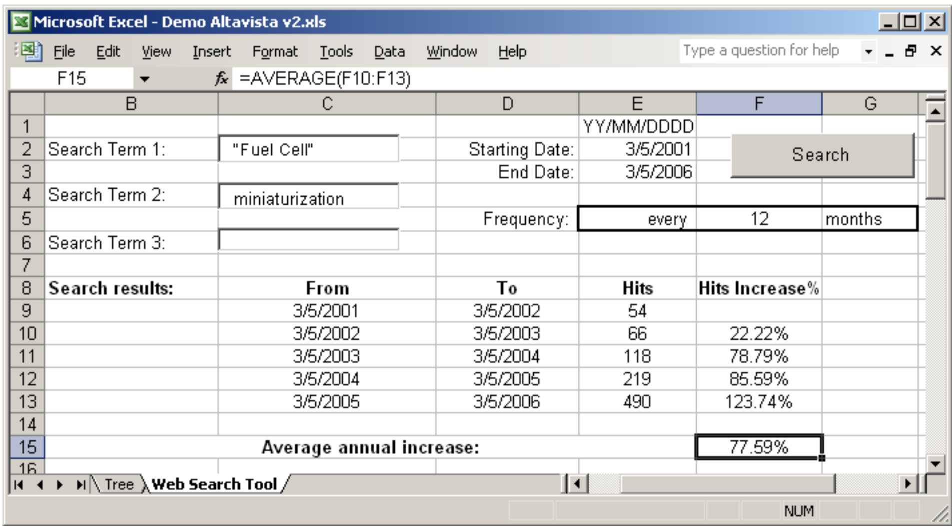


Figure 4: Experiment by searching





4. Calibration: Conditioning assessments using controls

The large and ever-increasing data on the web may mislead its users. A term that generates more hits by the agent does not necessarily imply that the event associated with the term is more likely to happen: it may well reflect no more than just the popularity of the term. More careful judgments must be applied when assessing the probabilities.

The field of information retrieval has been using techniques such as probabilistic search and vector models (Frakes and Baeza-Yates 1992; Salton and McGill 1983) to deal with similar problems. New enhancements to such techniques include incorporating advances in artificial intelligence and taking advantage of previous search experience by self or others (Chen 1995; Resnick et al. 1994). These methods can be highly complex and may require extensive human involvement, and hence may not suit our automated search environment. Moreover, it is beyond the scope of this paper to discuss in details the application of these methods to the topic at hand. Instead, we propose to use control searches to help experts calibrate probability assessments. That is, we compare the results to those for controls. We would need to do this at the time of assessment, and this should be feasible though not trivial. Here, the controls are other phenomena whose certainty is in some dispute. As a qualitative test, searches using the phenomenon name and the same adjectives (conclusive or inconclusive) were performed for phenomena ranging from certain (earth round) to ones now considered unlikely (vitamin E prevention), in table 1. Using this scale, we find that the statistic for global warming

seems to be in the lower end of the range for accepted (but not certain) phenomena. The expert could calibrate assessments using these facts.

Table 1: First query results.

Hits for phenomenon +	Vitamin E prevention	Silicone cancer	Homeo-pathic	**Global Warming**	WMD	Evolution	Earch round
Conclusive	2040	744	1550	**6140**	3270	63400	20600
Inconclusive	1860	525	1040	**3570**	1860	29100	8120
Ratio	**1.1**	**1.4**	**1.5**	**1.7**	**1.8**	**2.2**	**2.5**

Again, experts should consider empirical data about proposed measures when they provide conditional probability assessments.

The idea of calibrating the assessments can be extended. We start with results for analogous situations where the state of nature is known and, given the proportion of situations that lead to the different statistics, we use Bayes' rule to estimate the probability of the state of nature (for those situations) given the statistic. This only works when it is possible to consider a sufficient comparable population.

This is illustrated using an entirely different example. Suppose we wish to predict whether a given university is facing budget cuts. We can do a search on the name of the university and the term "budget cuts", and to make it more reliable, we duplicate the search substituting "budget increase" and use as our statistic the ratio of hits in the two cases. There are thousands of universities and colleges, so it is possible to do many "control" searches and find out for schools facing budget cuts what percentage have a hit ratio above 3, above 2, above 1, etc. This distribution is transformed into a distribution where we have percent of schools facing budget cuts as a function of the hit ratio. Very few schools with hit ratios at the low end, such as the University of Chicago, have faced budget cuts, while many or most of the schools with hit ratios at the high end (>2) have faced budget cuts. Queries were actually performed for only a small sample shown in table 2, and to do this in practice would require an automated query tool (script) to run the query using a long list of school names in sequence. It is necessary to check the relevance of the results before incorporating them in a model; a different search we conducted for corporation names in conjunction with the words "bullish" or "bearish" showed no correlation between stock performance and the ratio of favorable to unfavorable hits.

Table 2: Second query results.

Hits for school +	Chicago	Cornell	MIT	U of Ill	U Wisc	Georgia	U Oregon	South Carc	UMass
Budget + Cut		554	777						285
Budget + Increase		609	665						79
Ratio	**0.9**	**0.9**	**1.2**	**1.2**	**1.3**	**1.5**	**1.8**	**2.1**	**3.6**

Sensors could also perform purely randomized experiments to sample the contents of cyberspace. For example, in checking the proportion of web addresses that have been claimed, an agent could randomly generate to create 4-letter sequences, and then checked which ones were claimed as dot.com addresses, and the resulting percentage could be used as a metric. Alternatively, the agents could be directed to perform multi-stage actions, e.g., choose a second search based on the results of a first search (as humans do when they surf the web). In our example, we might first search for "new fuel cell technologies," then identify the most frequent company name (operationalized as a word not in the English language on the first page of hits, currently Enova), and finally search on that name and count hits on it (47,900) and use this last number as the statistic.

In this model, there is no reason at all that multiple agents cannot be used together. This lends still more flexibility in the design of the model. For example, predictions for future CAFÉ regulations may rely on both the monitor agent's findings from the IEM about the future presidential election, and the sensor agent's findings about the incidence of hits for global warming, etc.

Although not actually tested for the problem at hand, it is easy to generate an almost unlimited number of concepts for probes that operate on the open system of cyberspace. These could include tracking responses to inquisitive or intentionally provocative (predefined) comments posted to a (predefined) listserv or mailing list; offering an item for sale (e.g., on ebay.com) and tallying bids; similarly automated creation of a new predictive market for which price data will be tracked, a testing response speed for a site (a help site, an information site); counting visitors to a new website that is created and posted to Google using an automated script; tracking where visitors come from, or which links they select from such a site.

As an example, we created and tested a flypaper type of probe. This probe only affects automated programs, not live people. The probe consists of a new email account, created expressly for the purpose of conducting this study, and sending this email account name to a set of locations, some of which might respond with email announcements, advertisements or other contacts. The statistic of interest is the number of email messages the account receives within a given time period from when its name is submitted. It is necessary to preview the candidate targets, in order to prepare the necessary information for them. A test example chosen with the aim of receiving a sizeable response quickly (and without causing any actual expenses to be incurred by the provoked agents) was to register as a user at www.realtor.com, by filling out a form which requests demographic and geographic information about the user and where the

user must indicate areas of interest. This probe could, in the future, be used as a indicator of the level of one type economic activity in a given region (especially if the data entered and preferences selected were varied), and other distributions over variables could be assessed conditionally based on it

Many of these probes require prior preparation of material to be used in the future, and would also require us to define bounds of their action, including expense, time of operation, etc. With probes, along with significantly added power to obtain precise and unique real-time data, there are significant ethical concerns, obviously, because they can affect other people. There are also practical concerns resulting from the fact that once the system has been changed, the experiment cannot be run the same way again – an inquisitive or provocative comment posted to a specific listserv, for example, would have much different social meaning (becoming a nuisance) the second time it is posted.

Finally, because of their impact on the environment, these searches are perhaps more likely to have a financial cost, which, at some point could require an expanded set of agent capabilities, i.e., making decisions, executing financial transactions, and in conjunction with these two functions, performing value of information calculations.

5. Extension: When information is costly

The system designed by Tseng and Gmytrasiewicz (2002) calculated expected value of information (EVI) on the fly. For more flexible agents, a general methodology is needed to facilitate such calculations. A reasonable approach would be to manually assess expert probabilities for outcomes of costly experiments at the time the model is created, or before it is run. A refinement of this idea is to identify relevant free data or experiments, and actually include these as conditioning nodes to the nodes in which EVI is used.

Beyond the mostly technical aspects of integrating agents with decision analytic models, there is also a human process to consider. In traditional decision analysis, there is an analyst and a decision maker. The decision maker has the problem. The analyst talks to the decision maker (or designated subject matter experts), structures a model and obtains probability estimates, and then goes through the implications of the model with the decision maker. Incorporating web-based agents adds a level of complexity and there are new role for the people involved. As in standard DA, the *decision analyst* and the *decision maker* (or subject matter experts) create a basic influence diagram. Next, the decision maker lists potential information sources. An *information specialist* helps to identify corresponding web-based information, and conducts manual pilot tests to reduce the list based on where usable data may be found. The decision analyst and decision maker evaluate the relevance of courses in the decision model and the analyst determines what will have to be assessed and when. The information specialist and a *programmer* then script agents and connect them with a standard decision model (usually building

spreadsheets and/or other publicly available software). The decision analyst then may prepare for assessments by having the information specialist conduct trials that can be used as calibrating data, and finally assesses probabilities with the decision maker. When these probabilities are entered into the model, it is saved for future use.

Although we have considered basic building blocks for such systems, much more is required for them to be reliable. Certainly, strong error handling is necessary – there is no guarantee that designated sites will even exist in the future. When the data can be found, it should be validated to the extent possible, e.g., compared against upper and lower bounds beyond which it is not reliable. It may even be desirable to incorporate *paranoid* agents whose purpose is to seek red-flag type variables influence is not easily quantified (e.g., a marked increase in hits for the phrase "Armageddon" – currently yielding 2.2 million hits – beyond a pre-specified safe range) for the primary purpose of identifying major system changes. When confronted with missing data, the system might alert the user, or might be set to use the original assessment as a default. It will be important to track the age of data, as assessments where the conditioning information is old may become obsolete.

As mentioned before, there may be ethical (or even legal) concerns about some agent actions that system designers can imagine, and if this is a possibility, there should be guidelines and a vetting process. Finally, the system may be susceptible to mistakes resulting from changing societal frames over time. Data may not mean what we expect it to mean, e.g., the word "Bad" can mean bad or good depending on the social context.

6. Applications

When would this approach be worth the additional effort it requires? It is perhaps best to think in terms of decision classes and how this approach would be applied in each. In general, we would want to apply it when the model must be repopulated periodically or repeatedly, or where parts of the model would be recycled, i.e., when frequent evaluation is valuable but tedious. For one-time decisions, while it may still make sense to pull data from the Internet, we would probably do it all by hand and the need to define terms in a way amenable to assessment and to collection by automated agents would be minimal.

Portfolio resource allocation decisions: One highly successful application of business decision analysis has been portfolio management, where numerous projects competing for funds must be coordinated to achieve a variety of objectives. Valuation of individual projects and the portfolio depend on the importance of these objectives with respect to each project, the probability of success of each project (especially common in technology portfolios), the market demand for the different products, and the prospective costs of further development. When the number of projects exceeds perhaps a dozen, the resources required to manage a portfolio become significant. In larger organizations,

portfolios can grow to hundreds of projects or more. Then, we may use the same set of commands at least once for most projects in order to populate models with data as we manage the portfolio over time. A network of beacons might also prove useful, where assessments for one project are conditioned on status or decisions made on other projects. Here, project models could incorporate agents that update their associated beacons to reflect either new probability assessments or decisions. As in current portfolio management practice (Allen 2000), different "master" models can then slice the portfolio multiple ways to estimate its cash flows, requirements, NPVs, and portfolio level risk, etc.

Continuous decisions: Continuous decision variables are typical in optimization problems. Sometimes, formal algorithms are impractical and many alternatives must be evaluated in order to apply heuristic approaches, such as genetic algorithms. These approaches are not typically used in conjunction with decision analysis, because of the human-time required. If evaluation can be automated using Internet agents, such methods are practical for more problems.

Repeated and dynamic decisions: Operational decisions such as pricing and sourcing, contracting, and sales force or advertising allocations must be made repeatedly, for each new product or program, and periodically (e.g., annually or quarterly) for each program. The decision makers for these decisions rarely have the planning resources to perform detailed forecasts. Automated influence diagrams would enable these decisions to draw on the kind of information that these individuals would like to use to assess trends, perhaps more reliably than a central planning department could and at far lower cost, allowing application on more frequent decisions than otherwise would be feasible. There is twofold potential for application here: A single model can be used several times for one product or recycled for new products.

Dynamic situations, where the decisions themselves are not so difficult as gathering the relevant quantitative information, require monitoring of trends. For example spotting arbitrage opportunities and picking stocks, speculating on investments, evaluating real options, choosing daily resource allocations and pricing in markets allowing continual change all require estimates of supply and demand as well as other industry trends. The challenge is not to synthesize quantitative data, it is to collect enough relevant qualitative data and integrate it quantitatively.

Coordinated, decentralized decisions: Genetic algorithm type approaches requiring many evaluations are especially useful for finding coordinated combinations of acts (and with Internet data, especially in cyberspace) that get desired responses. For example, we may want to identify good websites, good website names, good product bundles (or positioning bundles), and good ways of segmenting customers. Similarly, we may wish to conduct a whole series of converging experiments in order to invest where demand is, where the competition isn't, and where the potential technology is promising. All of these would be measurable to some extent by taking web search frequencies.

Decisions that are best left decentralized (due to incentives or administrative cost) but that require coordination would benefit from using common modules with communication between them , and low-cost evaluation.

Some realistic application possibilities:

We now consider a pair of examples to suggest connections between the Internet and the decision model that could be useful in realistic situations. One possible application is in Pharmaceutical R&D. Here, the value of an investment in advancing a new compound depends on probability of success, probability of substitutes (from the same or other technical areas) probability of direct competition, other competitive marketing activity, market demand for the type of product, demographics, laws, regulatory environment and market size. A second business application would be a real-estate DSS advising a set of customers on whether and where to buy a house, the purchase decision is based on price attractiveness (based on sales and property properties such as square footage), labor market, housing trends, mortgage rates, economy, rental trends, comparative school and crime reports. On a different note, this kind of model may be useful in national security and intelligence applications. For example, a decision maker might be considering various investments to protect against gathering threats. The value of various preventive measures taken at various locations could depend on the prevalence of threats – searches could consider specific modes, specific threat origins, specific targets; beyond this, it may be ethically justified to develop probes that very actively solicit responses from people online in order to gauge the mood of various populations. Of course, an application such as this one would require official approval.

Because of the difficulties regarding reliability of applying current understanding to future conditions, it is easier to envision its use as an augment to human decision making and judgment than as a replacement – possibly as a device for identifying situations that require further human intervention.

7. Conclusion

Because of the way that influence diagrams are defined, it is natural to populate them with external data, and in particular statistics provided by web-based agents. The agents themselves are not required to be rational – they are merely the tools for conducting an experiment on the Internet universe – only the human user must be rational in making use of their findings. We explored some of the technical possibilities for agents ranging from mild mannered monitors to semantically sophisticated sensors to provocative probes, largely within the context of a realistic business decision problem. In developing proof-of-concept versions of some of these agents, we also uncovered challenges in their use. We envisioned the roles and process for creating these models as a variation on traditional decision analysis.

The analytics of this approach work cleanly and in particular, the potential for sensors appears to be rich. The potential for probes may be an even richer area for future research. Although there are limitations on the applicability of this approach, where it is applicable, it provides a low cost way to incorporate valuable information into decisions.

References

Allen, M., *Business Portfolio Management.* New York, Wiley, 2000.

Ariely, D., Lynch Jr, J.G., and Aparicio IV, M. "Learning by Collaborative and Individual-Based Recommendation Agents," *Journal of Consumer Psychology* (14:1/2), 2004, pp. 81-95.

Berners-Lee, T., Hendler, J., and Lassila, O. "The Semantic Web," *Scientific American* (284:5), 2001, pp. 34-43.

Berg, J., Rietz, T. “Prediction markets as decision support systems”, *Information Systems Frontiers*, (5:1), 2003, pp. 79-93.

Bhargava, H., Power, D., “Decision support systems and web technologies: A status report”, working paper, 2001.

Bjorneborn, L., and Ingwersen, P. "Toward a Basic Framework for Webometrics," *Journal of the American Society for Information Science & Technology* (55:14), 2004, pp. 1216-1227.

Bowman, D., and Narayandas, D. "Linking Customer Management Effort to Customer Profitability in Business Markets," *Journal of Marketing Research* (41:4), 2004, pp. 433-447.

Bradshaw, J.M. "An Introduction to Software Agents," in: *Software Agents*, J.M. Bradshaw (ed.), AAAI Press, 1997, pp. 1-48.

Chen, H. "Machine Learning for Information Retrieval: Neural Networks, Symbolic Learning, and Genetic Algorithms," *Journal of American Society for Information Science* (46:3), 1995, pp. 194-218.

Chen, H., Chau, M., and Zeng, D. "CI Spider: A Tool for Competitive Intelligence on the Web," *Decision Support Systems* (34:1), 2002, pp. 1-17.

Frakes, W.B., and Baeza-Yates, R. (eds.) *Information Retrieval: Data Structures & Algorithms*. Prentice Hall, New Jersey, 1992.

Furbush, D. "Program Trading and Price Movement: Evidence from the October 1987 Market Crash," *Financial Management* (18:3), 1989, pp. 68-83.

Hess, T., Rees, Loren, P. Rakes, T. “Using autonomous software agents to create the next generation of decision support systems”, *Decision Sciences* (31:1), 2000, pp. 1-31.

Howard, R., “Decision Analysis:Practice and Promise”, *Management Science* (34: 6), 1988, pp. 679-695.

Howard, R., “Knowledge Maps”, *Management Science* (35:8), 1989, pp. 903-922.

Keeney, R., *Value Focused Thinking*, Cambridge, MA, Harvard University Press, 1992.

Kowalski, G. *Information Retrieval Systems: Theory and Implementation* Kluwer Academic Publishers, 1997.

Mueller, J.P. *Mining Google Web Services: Building Applications with the Google Api* SYBEX Inc., Alameda, CA., 2004.

Ravinder, H., Kleinmuntz, D., Dyer, J., ‘The reliability of subjective probability assessments obtained through decomposition”, *Management Science* (34:2), 1988, pp. 186-199.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," CSCW 94, Chapel Hill, NC, USA, 1994, pp. 175-186.

Salton, G., and McGill, M. *Introduction to Modern Information Retrieval.* McGraw-Hill, 1983.

Tseng, C., Gmytrasiewicz, P., “Real Time Decision Support System for Portfolio Management”, *Proceedings of the 35th Hawaii International Conference on System Sciences*, HICSS-35. 2002.

Widing, M., Blodgett, D., Braun, M., Jusko, M., Keisler, J., Love, R., Robinson, G., “DOE Technology Information Management System Database Study Report”, *Argonne National Laboratory.* Technical Memorandum ANL/DIS/TM-26, 1994.

Wooldridge, M. & Jennings, N. (1995), "Intelligent Agents: Theory and Practice", The *Knowledge Engineering Review* (10:2), pp. 115-152.