Interpreting Clinical Data Using Concept Maps

Daniel E. Snyder

Science Application International Corporation
1321 Research Park Drive
Dayton, OH 45432

Introduction

This abstract describes the application of concept mapping for assembling information as a methodology for interpreting clinical data. This effort is intended to be a demonstration that concept mapping is a highly effective means of capturing the understanding of the problem domain associated with the expanding information available in today's clinical environments. The use of concept mapping in conjunction with a Macintosh-based software tool called the Interpreter allowed clinical information to be integrated and indexed using multiple classification categories and associated keyword lists.

Background

Concept mapping is intended to reduce the knowledge acquisition bottleneck often associated with expertise in complex domains. Since expertise is principally composed of procedural and perceptual knowledge, it is sometimes difficult for the expert to recall and verbalize. Additional problems arise when a knowledge engineer tries to access, combine, and assemble the knowledge into representational structures such as rules, frames, or scripts. In many cases, the acquired knowledge becomes more complex than the chosen structure can support. In order to avoid simply discarding knowledge that cannot be fit into the chosen representation, concept mapping provides a user-centered knowledge acquisition approach. It supports both the expert and the knowledge engineer as knowledge users in that they are both can be actively involved in the mapping process, using the concept map as a shared medium for communication. Concept mapping allows specialized knowledge in the forms of heuristics and problem solving methods to be explicitly associated by the knowledge users with static facts and general knowledge.

Concept mapping was shown to be an effective method for identifying user requirements, as it offered the opportunity to access the 'information heeded' and 'information remembered' depending upon the level of intrusion (i.e., probing) that occurred during the elicitation process. The "information heeded" was elicited as a pilot was allowed to "think aloud", with only a minimal amount of intrusion by the interviewer. The pilot was afforded the opportunity to spontaneously access information that related to the particular phase of a mission that was under consideration, without the imposition of the knowledge engineer's point of view (i.e., uninformed access). The concept mapping technique also offered the opportunity to elicit the 'information remembered' as the amount of intrusion was increased, as the pilot responded to access specific information (i.e., informed access). The ability to access these interrelated processes frequently led to the generation of a more highly differentiated concept map. (McNeese, Zaff, Peio, Snyder, Duncan, & McFarren, 1990)

The concept mapping technique supported by the Interpreter software was used to collect, assemble, and analyze information for evaluating pilot-vehicle interfaces for aircraft. Information from pilots, design engineers and human factors specialists was collected in the forms of traditional interview notes, interactive interviews using concept mapping, questionnaires, and anthropometric data taken from an existing helicopter cockpit. The Interpreter was used to enter the information as concept maps. A relational data base was then automatically derived from the graphical concept maps. From the data bases, the information was categorized into areas of interest for the evaluation. The Interpreter automatically indexed the information in terms of keywords in each of the categories. (Gomes, Lind, and Snyder, 1993)

Concept Mapping the AIM-94 Data

A primary benefit in using concept mapping is for visualizing data. The relationships between elements of data are evident in the connections between concept nodes in the concept maps, and the overall patterns of data in the maps. Because the Artificial Intelligence in Medicine-94 (AIM-94) data sets contained a substantial amount of data, it was not feasible to map each data item into individual concept nodes. Therefore, the nodes were replaced by concept tables which could include multiple lines of data. The tables were used to segment the data into blocks of data collected during a single week. In this way, weekly progress of the clinical trials could be readily visualized.

Ten samples from the Insulin Dependent Diabetes Mellitus (IDDM) data sets were randomly selected from the seventy data sets made available to AIM-94 participants. The ASCII data sets were segmented into weekly blocks of data using a word processor. The blocks were then entered into concept tables in the Interpreter software using the Macintosh Scrapbook feature. The transfer of the data sets into the Interpreter was accomplished in about 30 minutes per data set. See Figure 1 for an example of a concept map of IDDM data.

Interpreting the AIM-94 Data

The domain description provided insights into the relevant features of the data that would be of interest to investigators using the data. Those relevant features were used to define 10 categories for automatically encoding the data in the concept maps. The Interpreter allowed key items to be specified for each category.

The first three categories were the three types of insulin used in the studies. By examining the data in the concept maps for codes 33, 34, or 35 in the concept tables, the type of insulin being administered could be identified as regular insulin, neutral protamine Hagedorn (NPH), or Ultralente, respectively. The next five categories were ranges of Blood Glucose (BG) levels measurements either pre-meal (codes 58, 60, 62, or 64), post-meal (codes 59, 61, or 63), or unspecified (code 48 or 57). The normal range of BG was taken to be between 80 mg/ml and 150 mg/ml. Since the domain description extended this range to 200 mg/ml, a second category was used for BG between 150 mg/ml and 200 mg/ml. The other categories were selected to be BG less than 80 mg/ml, BG between 200 mg/ml and 300 mg/ml, and BG greater than 300 mg/ml. The final category which appeared to be prevalent in the data sets was hyperglycemic symptoms. The code 65 was used to identify an occurrence of these symptoms.

Once the categories were selected, the Interpreter automatically searched the concept maps and identified the specific data items for each category. The items could then be color-coded in the concept tables to aid in visualizing categories of data. The spreadsheet feature of the Interpreter also provided the number of occurrences of each instance of the categories for all of the concept maps as statistics associated with the data (number of insulin doses, the types of insulin used, etc.)

Results

A poster presentation of this study will be provided at the AIM-94 session on Interpreting Clinical Data at the AAAI Spring Symposium at Stanford University.

Data Set 1

04-21-1991 9:09 58 100
04-21-1991 9:09 33 009
04-21-1991 9:09 34 013
04-21-1991 17:08 62 119
04-21-1991 17:08 33 007
04-21-1991 22:51 48 123
04-22-1991 7:35 58 216
04-22-1991 7:35 33 010
04-22-1991 7:35 34 013
04-22-1991 13:40 33 002
04-22-1991 16:56 62 211
04-22-1991 16:56 33 007
04-23-1991 7:25 58 257
04-23-1991 7:25 33 011
04-23-1991 7:25 34 013
04-23-1991 17:25 62 129
04-23-1991 17:25 33 007
04-24-1991 7:52 58 239
04-24-1991 7:52 33 010
04-24-1991 7:52 34 014
04-24-1991 12:00 33 004
04-24-1991 17:10 62 129
04-24-1991 22:09 48 340
04-24-1991 22:09 33 005
04-25-1991 7:29 58 067
04-25-1991 7:29 33 009
04-25-1991 7:29 34 014
04-25-1991 12:49 33 004
04-25-1991 17:24 62 206
04-25-1991 17:24 33 007
04-25-1991 21:54 48 288
04-25-1991 21:54 33 002
04-26-1991 5:52 58 077
04-26-1991 5:52 33 009
04-26-1991 5:52 34 014
04-26-1991 13:06 33 005
04-26-1991 17:26 62 228
04-26-1991 17:26 33 007
04-27-1991 10:03 58 259
04-27-1991 10:03 33 010
04-27-1991 10:03 34 014
04-27-1991 17:20 62 256
04-27-1991 17:20 33 008

04-28-1991 8:42 58 109
04-28-1991 8:42 33 010
04-28-1991 8:42 34 014
04-28-1991 17:06 62 096
04-28-1991 17:06 33 007
04-28-1991 22:30 48 200
04-29-1991 7:39 58 128
04-29-1991 7:39 33 009
04-29-1991 7:39 34 014
04-29-1991 12:00 65 000
04-29-1991 13:38 60 192
04-29-1991 13:38 33 005
04-29-1991 16:50 62 263
04-29-1991 16:50 33 008
04-29-1991 22:28 48 081
04-30-1991 7:40 58 179
04-30-1991 7:40 33 010
04-30-1991 7:40 34 014
04-30-1991 12:55 60 088
04-30-1991 12:55 33 004
04-30-1991 17:40 62 185
04-30-1991 17:40 33 007
04-30-1991 22:00 65 000
04-30-1991 23:06 48 104
05-01-1991 7:38 58 086
05-01-1991 7:38 33 010
05-01-1991 7:38 34 014
05-01-1991 12:00 65 000
05-01-1991 12:14 60 060
05-01-1991 12:14 33 004
05-01-1991 17:23 33 007
05-01-1991 17:23 34 163
05-02-1991 7:39 58 147
05-02-1991 7:39 33 010
05-02-1991 7:39 34 013
05-02-1991 13:30 33 005
05-02-1991 13:30 65 000
05-02-1991 17:45 62 207
05-02-1991 17:45 33 007
05-03-1991 3:00 65 000
05-03-1991 7:48 58 305
05-03-1991 7:48 33 010
05-03-1991 7:48 34 014
05-03-1991 13:23 60 081
05-03-1991 13:23 33 005
05-03-1991 17:14 62 151
05-03-1991 17:14 33 007
05-03-1991 22:08 48 085
05-04-1991 6:50 58 133
05-04-1991 6:50 33 010
05-04-1991 6:50 34 014
05-04-1991 12:00 33 005
05-04-1991 18:00 33 007
05-04-1991 21:40 48 063

05-05-1991 2:00 65 000
05-05-1991 9:10 58 183
05-05-1991 9:10 33 010
05-05-1991 9:10 34 014
05-05-1991 12:00 33 005
05-05-1991 16:00 65 000
05-05-1991 18:00 33 007
05-05-1991 22:18 48 282
05-05-1991 22:18 33 002
05-06-1991 7:42 58 091
05-06-1991 7:42 33 010
05-07-1991 18:00 33 007
05-08-1991 7:43 58 251
05-08-1991 7:43 33 010
05-08-1991 7:43 34 014
05-08-1991 11:50 33 005
05-08-1991 17:30 62 135
05-08-1991 17:30 33 007
05-08-1991 22:45 48 129
05-09-1991 7:33 58 087
05-09-1991 7:33 33 009
05-09-1991 7:33 34 014
05-09-1991 12:00 33 004
05-09-1991 18:00 33 007
05-10-1991 7:40 58 255
05-10-1991 7:40 33 010
05-10-1991 7:40 34 014
05-10-1991 13:30 33 004
05-10-1991 14:15 65 000
05-10-1991 17:30 62 343
05-10-1991 17:30 33 008
05-10-1991 20:54 48 097
05-11-1991 8:30 58 136
05-11-1991 8:30 33 010
05-11-1991 8:30 34 014
05-11-1991 13:40 33 004
05-11-1991 17:36 62 082
05-11-1991 17:36 33 007
05-11-1991 22:02 48 110

05-12-1991 9:46 58 236
05-12-1991 9:46 33 010
05-12-1991 9:46 34 014
05-12-1991 12:00 33 004
05-12-1991 17:00 62 175
05-12-1991 17:00 33 007
05-13-1991 7:50 58 108
05-13-1991 7:50 33 010
05-13-1991 7:50 34 014
05-13-1991 12:00 33 005
05-13-1991 17:00 62 300
05-13-1991 17:00 33 008
05-13-1991 22:17 48 076
05-14-1991 7:36 58 065
05-14-1991 7:36 33 009
05-14-1991 7:36 34 014
05-14-1991 12:00 33 005
05-14-1991 12:30 65 000
05-14-1991 17:26 62 043
05-14-1991 17:26 33 007
05-14-1991 22:30 48 130
05-15-1991 7:45 58 183
05-15-1991 7:45 33 010
05-15-1991 7:45 34 014
05-15-1991 11:55 33 006
05-15-1991 17:09 62 182
05-15-1991 17:09 33 007
05-16-1991 7:30 58 117
05-16-1991 7:30 33 010
05-16-1991 7:30 34 014
05-16-1991 17:41 62 075
05-16-1991 17:41 33 007
05-16-1991 22:41 48 069
05-17-1991 7:35 58 076
05-17-1991 7:35 33 009
05-17-1991 7:35 34 014
05-17-1991 12:00 33 005
05-17-1991 17:17 62 201
05-17-1991 17:17 33 007
05-18-1991 8:55 58 244
05-18-1991 8:55 33 010
05-18-1991 8:55 34 014
05-18-1991 11:40 33 005
05-18-1991 16:45 62 104
05-18-1991 16:45 33 007

Figure 1. An Example Concept Map of the IDDM Data.

References

Gomes, M.E., Lind, S., & Snyder, D.E. (1993). A human factors evaluation using tools for automated knowledge engineering. NAECON '93 Proceedings, Dayton, Ohio.

McFarren, M.R. (1987). Using concept mapping to define problems and identify key kernels during the development of a decision support system. MS Thesis. AFIT/GST/ENS/ 87J-12 Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio.

McNeese, M.D., Zaff, B.S., Peio, K.J., Snyder, D.E., Duncan, J.C., & McFarren, M.R. (1990). An advanced knowledge and design acquisition methodology: application for the Pilot's Associate. AAMRL-TR-90-060. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Zaff, B.S., McNeese, M.D., & Snyder, D.E., (1993). Capturing multiple perspectives: A user-centered approach to knowledge and design acquisition. Knowledge Acquisition, 5, 1, 79-116.