Education

Editors: Gitta Domik and Scott Owen

Preparing Undergraduates for Visual Analytics

Ronald A. Rensink
University of British Columbia

Visual analytics (VA) combines the strengths of human and machine intelligence to enable the discovery of interesting patterns in challenging datasets.[1] Historically, most attention has been given to developing the machine component—for example, machine learning or the human-computer interface. However, it is also essential to develop the abilities of the analysts themselves, especially at the beginning of their careers.

For the past several years, we at the University of British Columbia (UBC)—with the support of The Boeing Company—have experimented with various ways of preparing undergraduate students for VA. Although inspired by the need to prepare students to become visual analysts, the result turned out to be fairly general in scope, applicable to other analytical approaches, as well as more general research. In hindsight, this makes considerable sense. Although the visual component of VA is necessary, it is insufficient; many analytical activities at the human end involve nonvisual skills, such as effective decision-making and the ability to quickly focus on the relevant parts of a problem.

The result of this experimentation is a third-year undergraduate course titled Cognitive Systems 303 (COGS 303) that focuses on "VA unplugged"—that is, on developing investigative abilities prior to training on the VA systems themselves. It was felt that if students focused on developing basic analytical habits of thought prior to learning VA systems, these habits would be reinforced by subsequent practice on "live" systems. As such, this course can serve as preparation for a more technically based VA undergrad or graduate course, such as that of Niklas Elmqvist and David Ebert.[2] Indeed, the course content could be considered an expansion of the initial few weeks spent by Elmqvist and Ebert's course on "analytical reasoning." Combining this with a more technically based course is the main way of teaching VA at UBC, and it is currently being considered as an element of various professional master's degree programs.

COGS 303

COGS 303 is a standard, three-credit, 12-week course open to any upper-level undergraduate in the constituent departments of UBC's interdisciplinary Cognitive Systems Program: computer science, linguistics, philosophy, and psychology. It has been offered—and continually revised—twice each year since the fall of 2008. Classes typically have about 30 students, although larger sizes could be possible. COGS 303 assumes a second-year background in statistics and a second-year course in human perception and cognition, but nothing beyond that. It is offered in the third year to balance two competing factors: students must be sufficiently mature to learn the analytic techniques, but young enough to apply them in their undergraduate career. Application of the analytic skills learned is typically via a subsequent directed-studies course done in a lab or nonprofit organization or, because many of the skills apply to research in general, subsequent courses on research work.

The exact form of the course has varied slightly over the years; Table 1 shows a typical syllabus. The main course textbook is *Asking the Right Questions* by M. Neil Browne and Stuart Keeley.[3] (See the "Additional Reading" sidebar for the other required materials.)

As is evident, the course has three sections, each focusing on a different level of skill. Each section has eight classes, which at a rate of two per week (at 80 minutes a class), require about a month to cover. These sections are crosscut by two recurring themes: the need for continual improvement and the need for effective communication. These themes would likely help in any course, but they are especially relevant in the training of analysts.

The course has been a favorite of most of the students who have taken it, receiving an average rating of 4.4 on a five-point scale. Of the students that eventually became analysts, all those who were asked reported that it played an essential role in their professional development.

Published by the IEEE Computer Society

Table 1. Typical syllabus for COGS 303.*

Date	Essay	Topic	Readings
0. Basics			
Week 1		Intro to the course, mindsets	Dweck
Section 1: Evaluating an Argument			
Week 2a	1-A	Belief versus knowledge	Burton; B&K, chap. 1
Week 2b	1-B	Structure of arguments	B&K, chap. 4; Booth et al.
Week 3a	–	Meanings of terms	B&K, chap. 5; Niederman & Boyum; Freedman
Week 3b	Debates I		
Week 4a	2-A	Assumptions	B&K, chap. 6; Heuer, chap. 6
Week 4b	2-B	Reasoning	Taleb; B&K, chap. 7
Week 5a	–	Evidence, recap	B&K, chap. 8, pp. 106–116
Week 5b	Debates II		
Section 2: Finding an Explanation			
Week 6a	3-A	Observation and mindset	Heuer, chap. 2; Beveridge, chap. 8
Week 6b	3-B	The role of hypothesis	Heuer, chap. 4; Beveridge, chap. 4
Week 7a	–	Intuition	Beveridge, chap. 6; Claxton
Week 7b	Debates III		
Week 8a	4-A	Imagination, visualization	Beveridge, chap. 5; Brown
Week 8b	4-B	The role of chance	Gilovich; Abelson, pp. 1–11
Week 9a	–	Choosing among alternatives, recap	Cadsby; Heuer, chap. 8
Week 9b	Debates IV		
Week 10a	Midterm exam		
Section 3: Systematization			
Week 10b	–	Experiment design	Cohen
Week 11a	T1-B	Comparative analysis	Yoon, Mayr, Kruskal & Wish
Week 11b	T1-A	Hierarchies, networks	Simon; Barabási, pp. 55–64
Week 12a	T2-B	Power laws, 80/20 rule	Barabási, pp. 65–73, 79–92
Week 12b	T2-A	Research questions, recap	Meltzoff, pp. 13–30; Abelson, pp. 11–14

** Tn = target essay n (an essay along with a presentation about the topic covered). A, B = cohort (the part of the class handing in an essay or presenting that day). B&K = Browne and Keeley textbook.[3] See the "Additional Reading" sidebar for the other required materials.*

Analytic Skills

The heart of COGS 303 is the development of useful analytic and research skills. This is done in three stages. Each is handled via a set of classes (or section) that aims to create a "layer" of particular skills, which is subsequently incorporated into later ones (see Table 1).

Most classes begin with a brief quiz on the readings assigned for that day. (Students are assumed to have done these at home.) The importance of the topic is then briefly discussed so as to provide further motivation. The remainder of the class is devoted largely to group work that exercises the skills covered. Most classes involve the analysis of some essay by each group; many include additional exercises for reinforcement.

Section 1: Evaluating an Argument

The first section focuses on critical thinking. Material is based on the Browne and Keeley textbook,[3] although other texts covering similar material could have been used. Topics include the following:

- *Belief versus knowledge.* The goal here is to shake the conviction of students that their beliefs are necessarily correct. Case studies and neurological reports show that the degree of intensity of a belief does not always correspond to the likelihood that it is true.
- *Structure of arguments.* Students are next shown that the way to higher-quality knowledge is via careful, structured argument. Material is presented concerning the structure of arguments (such as premises and conclusions, role of assumptions, and descriptive versus prescriptive arguments).
- *Meanings of terms.* Here, the goal is to show students that the meanings of words are not always

Additional Reading

The following list of materials details the sources for the required reading other than the main Browne and Keeley textbook.[3] The order is that in which they are encountered by students.

- C.S. Dweck, *Mindset: The New Psychology of Success*, Random House, 2006, pp. 3–11, 173–177.
- R.A. Burton, *On Being Certain*, St. Martin's Press, 2008, pp. 7–20.
- W.C. Booth, G.G. Colomb, and J.M. Williams, *The Craft of Research*, 2nd ed., Univ. of Chicago Press, 2003, pp. 114–123.
- D. Niederman and D. Boyum, *What the Numbers Say*, Broadway Books, 2003, pp. 60–63.
- D.H. Freedman, *Wrong: Why Experts Keep Failing Us—And How to Know When Not to Trust Them*, Little, Brown, & Co., 2010, pp. 37–41.
- N.N. Taleb, *Fooled by Randomness*, Random House, 2004, pp. 185–203.
- G. Claxton, *Hare Brain, Tortoise Mind*, Ecco Press, 1997, pp. 85–95.
- J.R. Brown, *The Laboratory of the Mind*, Routledge, 1991, pp. 1–7
- T. Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*, Free Press, 1991, pp. 9–28.
- R.P. Abelson,*Statistics as Principled Argument*, Erlbaum, 1995, pp. 1–11.
- T. Cadsby, *Closing the Mind Gap: Making Smarter Decisions in a Hypercomplex World*, BPS Books, 2014, pp. 88–98, 123.
- P.R. Cohen, *Empirical Methods for Artificial Intelligence*. MIT Press, 1995, pp. 1–10, 67–79.
- C.K. Yoon, *Naming Nature*, Norton, 2009, pp. 5–10.
- E. Mayr, *The Growth of Biological Thought*, Belknap, 1982, pp. 30–32.
- J.B. Kruskal and M. Wish, *Multidimensional Scaling*, Sage, 1978, pp. 7–16.
- H.A. Simon, *The Sciences of the Artificial*, 3rd ed., MIT Press, 1996, pp. 183–197.
- A.-L. Barabási, *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*, Penguin, 2002, pp. 55–73, 79–92.
- J. Meltzoff, *Critical Thinking About Research*, APA Press, 1997, pp. 3–30.
- R.P. Abelson, *Statistics as Principled Argument*, Erlbaum, 1995, pp. 11–14.

as clear-cut as usually believed—there is always some mismatch between word and reality. Emphasis is placed on the dangers of ambiguity and equivocation. Operationalization of terms is also covered.

- *Assumptions*. This introduces the various kinds of assumptions that can be made, such as descriptive assumptions (about facts) and prescriptive assumptions (about values). Students are shown that definitions can also contain assumptions.
- *Reasoning*. Students are shown that their "natural" thinking is not always rational—explicit training is essential. Also covered is the distinction between System 1 (reflexive thinking) and System 2 (effortful thinking) proposed by Daniel Kahneman and Amos Tversky.[4]
- *Evidence*. Students develop a feel for what is good evidence and what is bad, and they learn to avoid standards that vary according to personal beliefs. This includes the effects of personal bias, appeals to authority, the preference for the concrete over the abstract, and various fallacies concerning the believability of statistical studies.

Section 2: Finding an Explanation

Once a basic layer of critical thinking is in place, the next step is to develop skills for evidence-based investigation. Material is largely drawn from two texts. This first is *The Psychology of Intelligence Analysis* by Richards Heuer,[5] which describes the strengths and weaknesses of human cognition when applied to the intelligence domain; the second is *The Art of Scientific Investigation* by W.I.B. Beveridge,[6] which covers the skills used in scientific research, such as observation, intuition, and imagination. (In some ways, the content of COGS 303 could be viewed as a successor to Heuer's book, covering many of the same concerns, but drawing upon more recent developments in human cognition and involving a somewhat broader range of issues.) And in spite of the difference in the domains covered by the two books, there is a strong overlap of the skills discussed, which helps drive home the point that analytical techniques can be general in their applicability.

The topics in this section include the following:

- *Observation and mindset*. Students are made aware that perception is not passive, that what they observe is strongly influenced by what they expect (which in turn is based on what they know). Examples include the distortion of evidence to fit beliefs and the difficulty in changing one's mind so as to see new things.
- *The role of hypothesis*. As a follow up, the goal here is to show that hypotheses can help cope with incomplete information, but they must be used carefully. Discussion includes the strengths and weakness of three different kinds of hypotheses: situational analysis, theory, and analogy.
- *Intuition*. The goal here is to show that the nonconscious mind is a distinctive form of intelligence with particular strengths and weaknesses. (Explicit connections are made to the System 1 and System 2 distinction introduced earlier.) Discussion and exercises include how to effectively

coordinate the use of both kinds of intelligence.

- *Imagination, visualization*. Here students are shown the power of visual thinking, which is portrayed as a distinct form of intelligence with its own strengths and weaknesses. Applications discussed include the use of VA systems to discover new hypotheses. (The visualizations introduced here are also used in later classes when discussing relevant points.)
- *The role of chance*. The goal here is to have students realize that the patterns they detect are sometimes only the result of chance. Therefore, more objective tests (for example, those based on statistics) must be used before asserting conclusions with any confidence.
- *Choosing among alternatives*. Students are taught to make more effective decisions via the Analysis of Competing Hypothesis technique used by intelligence analysts.[5] This is broadened to include the case of several alternatives—not only competing hypotheses, but competing options in general.

Section 3: Systematization

After the basic analytical skills have been developed, attention turns to advanced strategic skills to guide higher-level thinking. This involves a rather heterogeneous set of topics; some focus on developing a feel for complex systems, while others are concerned with more conventional topics such as experiment design. In accord with its heterogeneity, the material for this section is drawn from a variety of sources (see Table 1 and the sidebar). The following topics are covered:

- *Experiment design*. The goal here is to give students a feel for the kind of data to obtain when data either isn't available or is unsuitable (for example, the need to establish causation rather than just correlation). This section also includes discussion of exploratory versus confirmatory analysis.
- *Comparative analysis*. Students are introduced to various ways of investigating complex systems when a controlled experiment is not possible. Examples include natural experimentation (where variation occurs naturally) and classification (where correlations across several dimensions create groups).
- *Hierarchies, networks*. The goal here is to make students aware that some nonlinear systems can have an interesting structure (such as hierarchies) and can sometimes have counterintuitive properties (such as small-world connections).
- *Power laws, 80/20 rule*. Students are introduced to another important characteristic of interacting systems: power-law distributions of properties. This includes a discussion of how to use such distributions (when they exist) to become effective in various activities.
- *Research questions*. The focus here is on finding a question for which the answer will make an important difference. Although of primary interest to researchers, this topic is also somewhat relevant to analysts, if only to help determine which issues to consider.

Crosscutting Themes

In addition to analytic skills, two other kinds of skills are developed: those that enable the student to continually improve their performance (including their intelligence) and those involving communication. These are handled via two crosscutting themes, usually covered via the exercises carried out in each class.

Theme 1: Continual Improvement

This theme is largely drawn from the work of Carol Dweck,[7] who argues that to learn more effectively students must move away from viewing intelligence as a static, fixed quantity (like eye color) and begin to view it as a dynamic quantity that can increase with practice (like strength). Students are encouraged to "play around" with problems and not overly worry about making errors or looking silly—the goal is not to avoid mistakes at all costs, but to learn from them.

The theme of continual improvement is sufficiently important that much of the introductory week is devoted to it. Much of the information that students use for improvement is obtained via peer evaluation in class; such peer-instruction techniques have been shown to help students consolidate new material and apply it to concrete situations.[8]

Theme 2: Effective Communication

The second theme concerns communication skills. These are essential to analysts, if only to enable them to clearly understand the material they encounter and to report their results clearly and concisely. Two forms of communication are emphasized:

Writing. Students are asked to write and then analyze numerous essays, with the topic being the focus of the class that day. A 10-minute session on writing is given at the end of the first few classes, based on the classic grammar book by Strunk & White.[9] Each student is asked to write six essays, each of no more than 400 words, generally at the

rate of one every other week. The essay topic usually involves the material under consideration for that class. The limit of 400 words is severe, but it motivates students to be concise.

Analysis is done by groups of four to five people, and the goal is to have students learn from their peers, allowing best practices to spread. Each group is given three copies of an essay from a classmate and then asked to determine its two biggest weaknesses, in terms of both content and style, and its main strength. Each essay (and group analysis) is marked by an instructor or a teaching assistant. Although only a fraction of the essays submitted can be analyzed by a group during a given class, an effort is made to ensure that each student gets feedback from at least a few groups over the duration of the course.

In accord with the theme of continual improvement, groups are also asked to suggest improvements for the problematic parts. Borrowing a technique commonly used in writers' workshops, the essays have no names attached—instead, only a nonpersonal identifier is used (such as the name of a superhero) known only to the student and the instructors. Thus, groups need not deal with any awkward issues that might arise from analyzing the work of someone they may know. On occasion, a group may receive an essay written by one of its own members. Students are asked not to give away their identities so they can hear what others really think of their work. Students often report this to be a rather educational moment.

Verbal presentation. To help develop this skill (as well as practice their analytic skills), teams of two to three students occasionally debate each other on a selected topic. To encourage conciseness, each student must give their address (opening statement, rebuttal, or summary) in less than three minutes. Participants are given written feedback by the other students regarding content (for example, how well the topic was addressed) and style (for example, how clear the presentations were). To encourage mental flexibility, a team is not assigned its position on the given issue—pro or con—until a coin toss just before the debate begins.

Once students have done a few debates, they are asked to give individual presentations about their target essays. This involves a verbal delivery along with a set of slides. As in the debates, the audience gives written feedback. To encourage students to relax, this feedback is qualitative only (for example, whether the student spoke too quickly) and does not affect their mark, which derives from the essay they wrote. Presentations are again limited to three minutes, in part to allow several students to be assessed in each class and in part to encourage people to be brief. At the end of each presentation, the class is asked to give the presenter a standing ovation, no matter what, to honor the effort made.

The modular design of COGS 303 allows it to be adapted in various ways. For example, if students are already familiar with critical thinking, the first section could easily be dropped and more emphasis could be placed on other topics (such as separate classes for imagination and visualization or more time on logical fallacies). If time is limited, section 2 (along with a few classes from section 3) could be inserted into a more technically based course to provide a nice combination of basic VA and analytic skills.

For more information about this course, visit the COGS 303 website at www.cogsys.ubc.ca/303/.

References

1. J.J. Thomas and K.A. Cook, eds., *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE CS, 2005.
2. N. Elmqvist and D.S. Ebert, "Leveraging Multidisciplinarity in a Visual Analytics Graduate Course," *IEEE Computer Graphics & Applications*, vol. 32, no. 3, 2012, pp. 84–87.
3. M.N. Browne and S.M. Keeley, *Asking the Right Questions: A Guide to Critical Thinking*, 11th ed., Pearson Education, 2013.
4. D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2013.
5. R. Heuer, *The Psychology of Intelligence Analysis*, Center for the Study of Intelligence, 1999.
6. W.I.B. Beveridge, *The Art of Scientific Investigation*, Blackburn Press, 2004. (Originally published 1957.)
7. C.S. Dweck, *Mindset: The New Psychology of Success*, Random House, 2006.
8. A.P. Fagen, C.H. Crouch, and E. Mazur, "Peer Instruction: Results from a Range of Classrooms," *The Physics Teacher*, vol. 40, no. 4, 2002, pp. 206–209.
9. W. Strunk Jr. and E.B. White, *The Elements of Style*, 4th ed., Longman, 1999.

Ronald A. Rensink *is an associate professor in the Departments of Computer Science and Psychology at the University of British Columbia. Contact him at rensink@cs.ubc.ca.*

Contact department editor Gitta Domik at domik@uni-paderborn.de and department editor Scott Owen at sowen@gsu.edu.