Tidbits from AAAI’96
by Ramesh Patil, Associate Program Chair, AAAI '96

Many of you participated in the organization of the upcoming AAAI National Conference either as an author or as a reviewer. I was privileged (or stupid enough to volunteer!) to be associate program chair responsible for the reviewer self-selection process. In this column, I would like to share my experiences and observations while they are still fresh in my mind.

In other conferences, reviewers are assigned papers by the conference program committee. As a result, reviewers are sometimes assigned papers on subjects that they are not very familiar with. Instead, reviewers for AAAI submissions can suggest to the program committee the papers they would like to review. This idea of self-selection was initially proposed by Matt Ginsberg during the 1991 program-committee meeting. The 1992 program chairs went along with the idea with a great deal of skepticism, but couldn't refuse as Matt volunteered to do all the work needed for self-selection. In 1992, AAAI generated a list of titles, keywords, and authors from the papers received. Matt sent them out by email, and the reviewers returned their selections back to Matt by email. Matt spent considerable time collating the results and presenting them to the program-chairs to aid in the assignment of reviewers to papers. Clearly, the program chairs found that the self-selection data greatly simplified their task. Individual reviewers (who participated) found that (continued on page 7)

Hans Chalupsky
Wins Newsletter Name Contest

We are pleased to announce that Hans Chalupsky has won the ISD Name the Newsletter contest. Hans’ submission, “The InSiDer” received the most first and second-place votes from the 32 ISDers who voted. He will receive a $20.00 gift certificate from California Pizza Kitchen for his winning entry.

Thanks to all of the creative people who entered the contest!

ISD Logo Contest

Put on those creative thinking caps, we’ve got another contest! Bill Swartout would like to encourage the ISD staff to create an ISD logo.

The only rule is that the logo must be ISD related.

To enter, submit a hard copy of your logo to Theresa Cox by April 26. We will display the logos for the division to vote on. The contest winner will receive a gift certificate and will be announced in the next newsletter.

Inside This Issue

1 Tidbits from AAAI ‘96
2 Project Overviews
3 New Faces at ISD
4 Getting to Know ISD
5 ISD Tidbits
IBM Pays Natural Language Group Large Pots of Money
Leonard D. Duck, Staff

The SENSUS Ontology, also known as the Pangloss Ontology, seems to be interesting to more than the research world. IBM Research Laboratories in San Jose acquired a research license for the Ontology early this year for a modest sum (the "large pots of money" was just to draw your attention).

This modest sum will either go toward buying a yacht for the NL researchers or toward further work on the Ontology (we are still debating which, since Matthew Haines already has a yacht in the Marina on which we hold wild parties every second full moon). We choose the latter option, then the work will involve bringing into alignment the Ontology Base (the top 250 or so concepts; also known as the Penman Upper Model) and the Upper Ontology from CYC [Lenat and Guha]. (For some reason, IBM is interested in the Upper Ontology of CYC. Maybe after looking at it for a month, we will be too. Or maybe we will be dead. Life takes strange turns.)

Fragments of the Ontology will be sold in Richard Whitney's office every afternoon from 4:30 to 5:00 pm, at $2.50 per 500 grams. Bring your own paper bag, and rush before IBM buys everything in stock!

The PROBES Project
by Lewis Johnson

PROBES is developing analysis and explanation capabilities for training simulations, in order to make such simulations more effective as training tools. It is part of the Exercise Management effort within DARPA's Computer-Aided Education and Training Initiative.

Military training simulations typically incorporate large numbers of entities, both human and computer-generated, which interact in a simulated world. Setting up such simulations is quite laborious, and once the simulation is running it must be monitored closely in order to assess whether the simulation is providing the human participants with the intended training experiences, and whether these experiences are having the desired effect. PROBES brings plan recognition and explanation technology to bear on this problem. Assuming that the training and mission objectives have been specified as part of the exercise scenario, PROBES observes entity behavior and recognizes behavior patterns (plans) that are relevant to those training objectives. It also provides a dialog interface to entities, enabling trainees to interact with the simulated scenario participants in "after action reviews." Trainees can find out from the participants why they did what they did. Integrating the plan recognition and explanation capabilities yields "talking probes," i.e., intelligent agents that can monitor activity in the simulation and provide the trainee with narrations and critiques.
Complex software is difficult to modify in part because it is difficult to understand; the greater the complexity the greater the understanding problem. Current work on the Dynamic Documentation project at USC/Information Sciences Institute has been addressing the software understanding problem through the development of software explanation technology, i.e., the dynamic generation of presentations specific to individual users' tasks and needs. Software explanations reduce the amount of time spent sifting through irrelevant information, helping to ensure that the presented information is more readily understood.

Software explanation technology has been focused to date on generation of textual or hyper-textual descriptions. This means that they have not been able to take advantage of other presentation media, such as diagrams and animations. The aim of the work proposed here is to develop software explanations that integrate diagrams, animations, and text. An enhanced tool called Media-Doc will be built that generates explanations that are contextualized and task-specific just as the current text-based software explanations are.

The proposed work will offer several significant innovations over existing graphically oriented software tools (e.g., CASE tools and visual languages). First, the generated presentations will be designed to meet specific communication objectives, highlighting interesting system features and suppressing repetitive details. They will explicitly employ graphical presentation idioms used by expert graphical designers to convey information effectively. They will seek to maintain a consistent level of complexity, regardless of the complexity of the software being explained, by limiting the amount of detail that is presented at any one time. We will also explore ways of making better use of spatial metaphors to help users to orient themselves better when viewing abstract diagrammatic presentations of software. Media-Doc presentations in contrast will have tightly integrated text, diagrams, and animation. The presentations will be interactive, enabling users to obtain additional information and clarifications of the information presented. Effective software explanation requires knowledge of the software's architecture and design rationale. This knowledge must be kept up to date as the software evolves. Media-Doc will therefore leverage work on program understanding and transformation in order to extract the necessary representations, e.g. by suppressing detail in complex designs in order to uncover meaningful abstractions. Media-Doc presentations will be editable, allowing users to annotate, critique, and modify designs directly via the presentation interface.

Collaboration with SYSTRAN on Machine Translation Funded

by Ed Hovy

The Office of Research and Development of the CIA has funded a one-year research collaboration between SYSTRAN from San Diego and the Natural Language Group. SYSTRAN is the world's oldest continuously operational machine translation company, selling MT services and systems in numerous languages, including English, Spanish, French, Russian, German, and Japanese.

Work, which starts this summer, will focus on the transition of technology and knowledge from ISI to SYSTRAN. The determiner insertion module from Japangloss (the module that decides when and where to put in the articles "the" and "a", which Japanese lacks) will be incorporated into SYSTRAN's Japanese-to-English system. This work, which is scheduled to take three months, will be followed by the development and implementation of a new paragraph-length output generation controller for SYSTRAN, in order to improve the output quality in places where English and Japanese sentence boundaries fall differently. Primarily involved from ISI will be Kevin Knight and Kenji Yamada (first project) and Eduard Hovy and Richard Whitney (second project).

Limerick

There once was a man from the sticks
With a passion to write limericks.
But he failed at the sport,
'Cause he wrote them too short.
Besides school and work, my interests include racquetball (which I try to play on a regular basis), ballroom dancing, hang gliding, skydiving, river rafting, mountain biking, and reading. My personal philosophy is that life is meant to be enjoyed.

Getting to Know ISD
In an effort to get to know each other better we will be highlighting members of the ISD staff in each newsletter. This month we are highlighting the administrative staff.

Sheila Coyazo
I was born in Waterloo, Iowa, but I didn't stay there long... :-) I've lived on the west coast since I was three, first in Portland, Oregon, for a couple of years, then in the San Francisco Bay area until I went to college, and in Southern California since 1978. My heart is still in the Pacific Northwest, and I make no secret of my ambition to someday drag my husband kicking and screaming to live in Seattle. (He was born in Yucaipa, CA, and hasn't ever lived outside Southern California.) Wish me luck...I'll probably retire here.

I began working on my BA in 1972 at UC Santa Barbara, when it was a haven for hippies and anti-war activists. By the time I dropped out in 1975, UCSB had reverted back to its "party school" image, and in my opinion it wasn't half as interesting a place to live. I spent the next three years working for the company that puts on the Renaissance Pleasure Faire and (frequently) collecting unemployment insurance or working temp jobs. Realizing that the real world wasn't so fun for people without degrees, I scurried back to college in 1978 to finish my degree (in history) at UC Riverside. While at UCR I met my husband, Joe, and rekindled my love of choral music and singing. Joe is a tenor and I'm a soprano, and since we met we've sung together in many choirs (we do weddings, too), most notably the Los Angeles Master Chorale. My greatest choral achievement was performing Beethoven's 9th symphony at the Hollywood Bowl with the Master Chorale and the LA Philharmonic, conducted by Carlo Maria Giulini. I could've died and gone to heaven at that point, my life was complete! :-) Joe is actually an opera singer disguised as an elementary school teacher. He has a master's degree in vocal
I have been happily married for almost 12 years to Jerry. And as most of you know I have two great boys (Mackey, 8, and Dylan, 1) who keep me extremely busy!

I like to do things which take me outside, like biking, camping, hiking and running (of course).

**Velda Thomas**

Velda is a native of Los Angeles, who lives with her mom, Lynda, near campus. She will be receiving B.A.’s in Psychology and Creative Writing from USC in May and first worked at USC/ISI as a work-study. She’s now a permanent fixture. This Fall she will enter the Professional Writing Program at USC, where she hopes to perfect her craft. She’s also looking forward to the publication of one of her poems in the upcoming anthology *Carvings in Stone*, due out this summer. This poet/singer hopes to learn how to sew and prepare a gourmet meal in the not too distant future.

**Theresa Cox**

I’ve lived all my life in the South Bay. While still in high school I worked at a nursery school. I love children (probably because I don’t have any), taught the 2-3 year olds and worked my way up to Assistant Director. I left to go to school at Brigham Young University. I was only there for two semesters as my asthma and those mountains just didn’t get along! When I returned to California I got a job at TRW and worked on the financial side of proposals, costing them, preparing budgets and putting the data together for the project leaders. I was given a pink slip in March, 1988. At the time I was devastated but can look back at that experience as a blessing in disguise. I landed a job with an Insurance Consultant, a decision I still can’t figure out. It was a small company, I would be the second employee, and the owner was an entrepreneur. While with this company I discovered my love for computers! We ordered an AT&T 286 with a 20MB hard drive and 1 meg of Ram! I thought I was in heaven (probably because I didn’t know any better) About a year and a half later our beach cottage office burned down (that’s another story) and I got to buy a new machine. The only parameters I had from my boss was that it had to be a true IBM (he was obviously computer illiterate) I bought a 386 with color monitor and thought I had died and gone to heaven.
Two years later we hired another employee. As the company grew I became the in-house system manager, software expert and computer fixer. I loved this part of my job and could have spent all day, every day working on the computers. Unfortunately my boss didn’t understand the time I would spend trying to solve a problem or fix a bug or why I felt we should be networked. He only understood that we in the insurance consulting business, not the computer business.

My background is in the financial side of business and I love numbers (would like a few more on my pay check!). Before I discovered computers I wanted to be a controller of a large corporation (someone like Cary Thomas). Now I find enjoyment in combining my love for numbers, spreadsheets and budgets with the fun of computers (although I wouldn’t call learning the Mac “fun”). I’m lucky enough to indulge both interests in my position as Administrative Coordinator at ISI.

I spend my “off” time going to school, where I’m working on a Business Finance major, Computer Science minor. I’m active in teaching quilt making. I also demonstrate a software program that allows quilters to design and create their quilts on the computer. I’ve worked with the developer of the software, making suggestions, beta testing, etc. I hope to be more involved in the actual programming in future versions. I’m also involved in a company that raises money for charities and non-profit organizations.

I also stay very busy as a homeowner. I sit on the board of directors for our homeowners association and head up our neighborhood watch program. This experience is teaching me a great deal about small town politics (something I’m not sure I want to know about!). I’m also responsible for the bi-monthly newsletter for the association.

Some other things that keep me busy are my nieces and nephews, a 1973 Chevy Nova that I like to tinker with (I keep threatening to sell it, but don’t seem to get an ad in the paper!) and my nephew’s baseball team which I help coach. I have an annual pass to Disneyland and go every chance I get. I spend 1-2 nights each month at the Genealogy Library researching my family history. I love softball, snow and water skiing, fishing, reading and tinkering around the house.

---

ISD Tidbits

Trustees Visit ISI
USC’s Board of Trustees recently visited ISI. Paul Rosenbloom and Kevin Knight presented the Soar and PANGLOSS projects. The presentation went very well, and one board member was overheard saying, “ISI is USC’s best kept secret!”

Welcome Jay Zhang!
The Zhang family has grown with the addition of their little boy, Jay Zhang who arrived on February 22. Jay weighed in at 9 pounds, 8 ounces and was 19 inches long. Weixiong Zhang and Hui Zhou are the proud parents. Jay has an older brother, Kevin, age 3.

Wayne would like to express his appreciation to Kary Lau, Bill Swartout, Wei-Min Shen, Lisa Moses, Ping Luo, Peter Will, and especially to Ramesh Patil, Paul Rosenbloom and their wives for their concern and help during the difficult time the Zhang family has been through.

GRA Engaged
Kittens Mough and Chutney are proud to announce the engagement of their owners, I-Doc project member, graduate student Amy Biemann and Scott Hughes, graduate student in Physics at the California Institute of Technology. Much to the chagrin of their parents, the couple has not even considered wedding plans yet. Thus, no date has been set. When asked for comment, the kittens said “Meow.”

A Little Story
Author Unknown
This is a story about four people named Everybody, Somebody, Anybody and Nobody. There was an important job to be done and Everybody was sure that Somebody would do it. Anybody could do it, but Nobody did it. Somebody got angry about that because it was Everybody’s job. Everybody thought Anybody could do it, but Nobody realized that Everybody wouldn’t do it. It ended up that Everybody blamed Somebody when Nobody did what Anybody could have done.
Happy Birthday!

Would you like to have a monthly ISD birthday celebration? For example we could choose the 3rd Friday of each month to celebrate whoever has a birthday that month in our division. Participants would donate $2 a month for cake, drinks and cards. We’re taking an informal survey to see how many people are interested in doing this. Let me know by e-mail <kary>.

Soccer Anyone?

Every Tuesday and Friday from noon to 1:30, a group of ISD-ers and other ISI-ers play pick-up soccer at Westchester Park. We meet at noon in the 9th floor elevator lobby, and carpool to the park. Come and join us. Players of all skills are welcome. Don’t miss the fine spectacle of age-old rivalries such as Ed’s and Ping’s one on one. If interested, send mail to szekely@isi.edu.

AI Seminars

April 1 Joost Breuker, University of Amsterdam "Structure Of and Access To the CommonKADS Library of Problem Solving Methods"

April 12 Richard Korf, UCLA "Automatic Discovery of Admissible Heuristic Functions"

April 23 Jaime Carbonell, CMU Title TBA

May 10 James Lester, North Carolina State University "Generating Natural Language Explanations from Large-Scale Knowledge Bases"

June 7 Martin Greenberger, "The Future of Computing and the Media: Travels into the 21st Century" NOTE: This talk is part of the Jacob Marschak Interdisciplinary Colloquium on Mathematics in the Behavioral Sciences at UCLA, and will take place at 1PM in room C-301 of the Anderson School Building at UCLA.

The days and nights of parenthood are very long, but the years are very short. 

T. T. Nhu

AAA'I '96

continued from front page

the papers they received for reviewing were more to their liking than in previous years. Although the implementation was primitive, the idea was clearly useful. Few thanked Matt for his work, and everyone gave him suggestions for improving the process. I had been one of Matt’s vocal supporters from the beginning, and thus fell to me the task of supporting self-selection the following year. Based on the comments received and my own experience, I had come to the conclusion that providing online abstracts to the reviewers would significantly improve the quality of data collected. But how was this to be achieved? Clearly, email could not handle this volume of data. The key challenge was to develop a solution that would allow us to capture electronic abstracts from authors without any special software, and to develop tools to aid reviewers in the process of self-selection that would be robust and portable enough to run at approximately 100 sites. In the pre-Web era, the clear solution for the first was to use plain ASCII abstracts submitted via email, with keywords marking each logical segment of the abstract. The solution for the second problem was unclear. After much thought, I decided that the greatest degree of portability could be achieved by using EMACS as a nearly universal platform, FTP as a means of sending data to reviewers, and email as a means of receiving the selection. But EMACS did not run on Macs or PCs. For the Mac users, I chose Mac Common Lisp (it’s for an AI conference, after all), and decided to ignore the PC users (how times have changed!). I spent about 2 months over a 6-month period (weekends, after-hours, vacation-time) developing the software, documenting it, and describing the self-selection process. I had the test version completed and tested by program committee members and felt ready. Boy, was I wrong!!

When the abstracts began arriving, it was clear that AAAI authors (scholars though they might be) do not read instructions, and even when they do, do not follow them. A great many abstracts did not have appropriate markers such as "TITLE:" to allow my scripts to parse them. Others added or removed blank lines, changed cases, centered, indented, LaTexed, and so on!! As the abstracts flowed in, I constantly extended my parser. to take care of as many
cases as possible, or edited the abstracts to make them acceptable to the parser. You would not believe how many people submitted LaTeX files (they are ASCII, aren't they?). Working fast and furiously, I managed to get all the data organized, only to find that the electronic abstract titles and the paper titles (entered by the AAAI data entry staff) did not correspond well enough to allow automatic merging of the two databases!! Another GOTCHA to remember for the future. Well, three days after the paper submission deadline, the self-selection message went out.

About 60% of the reviewers loved it (the rest either didn't care or were clueless). I received many compliments and questions about advanced features that I had added but had not documented (There are true hackers, even at very senior levels). However, I was surprised by the number of people who in today's politically correct terms can only be described as "computer challenged". Just to give examples, one reviewer requested line-by-line instructions on how to anonymously FTP a file, another requested that I interact with an on-site system expert at his site. I also discovered that EMACS versions are not all compatible, firewalls at many companies (IBM, AT&T) do not allow FTP, and on and on. Anyway, after a week of constant attention, the process was over! I collected and processed reviewer selections, and automatically generated the first-pass assignments for reviewers and a schedule for the program committee meetings. As a result of all the hard work, the paper-assignment process, which used to take up to three days before, was accomplished in less than a day, with superior assignment, fewer returned papers, and better overall reviewer satisfaction.

During the program committee meeting, Barbara Hayes-Roth (AAAI'94 co-chair), Howard Shrobe (conference chair), and I were standing around discussing self-selection (we didn't have papers to review!), when Barbara asked me if I would assist her with the process. When I refused, Howard stepped in and volunteered. "How hard could it be?" he said, even though I urged him not to, and warned him about the amount of work needed. I told him he was making a mistake, but I offered my help anyway — what are friends for?

After some thought and discussion, Howard decided to base the 94 self-selection on the software he and John Mallory were developing for the '94 presidential election. To summarize, the process of abstract collection was significantly better than in 93, but the self-selection was a disaster! Not because the software was bad, but because the reviewers were not adequately trained to use it (and it rejected anything that was not formatted right!), and the upper mid-west hub of the Internet went down (due to severe weather) during the critical period. We survived, much bloodied but wiser.

Luckily, 1995 was an IJCAI year — no AAAI, no self-selection. I was enjoying my retirement, having been elected a fellow of AAAI in the meanwhile, when Dan Weld called me. He had been asked to co-chair the '96 conference, but would do so only if I agreed to assist him in self-selection. After much convincing I agreed. Things had changed considerably in the last few years with the onset of the WWW. I believed that it could be done better — the challenge was there, and I was sucked into it yet again.

This time the software again took about 2 man-months of work spread over a year. In the process I have become an expert in Perl, CGI, HTML, web-browsers, and how to program for the least common denominator. Many of you have used this software, so I will not describe it here (Try URL "http://indra:8001/ncai97" for a demo), except for highlighting the key features. First, the form-based interface for abstract submission gave us a much more controlled environment for data input, and allowed us to dynamically allocate tracking ID's to abstracts and give that information to authors for inclusion with the papers. We had virtually 100% compliance from the authors, although some came in after the deadline. This eliminated the need for database merging. More significantly, this eliminated the need for data-entry at the AAAI office, consequently eliminating the need to hire temporary data-entry staff (saved $$$). The self-selection process on the Web was used by about 200 (out of 230) reviewers. We were able to collect more detailed information about user preferences and conflicts.

While abstract submission and self-selection were going on, the AAAI office and the program chairs were able to monitor the progress. The reviewer assignment process was considerably simplified. Had all reviewers
participated, we could have completed the assignment with half a day's work. More important, because the field has grown considerably, and the reviewer load has dropped to less than half that of previous years, many of the reviewers were unknown to us. Without self-selection and automatic reviewer assignment, it would have been impossible to maintain or improve on the high quality of reviewer assignment, and thus the high standard of reviewing that we have come to expect from the MAl conference. The most amazing anecdote about this year's process was that during the peak period of both the abstract submission and the self-selection processes, I was out of town. Yet the software worked without a hitch, and I was able to maintain the site, add last minute reviewers, etc., all over the Web!

As a side note: Now that the new Web browsers hide much of the complexity of using the net, we no longer have to deal with FTP challenged users, but we now have to deal with the Web-browser challenged!

An interesting side-effect of my experience with self-selection is that now I have three years' worth of on-line abstracts, a resource unique to ISD for information retrieval, data-mining and web-assistant research. If you have an interesting idea and would like to collaborate, please drop by. You may even get to play in the year's self-selection.

All this is fine and dandy. But how did we do? Well, we did very well. ISD had 15 submissions, of which 8 were accepts (53.3%). Overall 703 electronic abstracts and 634 paper submissions were received, from which 196 papers were accepted (31%).

What were the hot areas? This question can be answered by looking at the conference submissions if we assume that submissions in each area reflect the level of activities in that area at least for those areas that do not have specialized national conferences (e.g., machine learning, uncertainty reasoning, vision, robotics etc.). The following table gives the number of times each keyword appears in the 703 electronic abstracts.

If you would like additional information about the conference, papers submitted, or self-selection, but most importantly, if you would like to participate in supporting self-selection next year, please drop by for a chat.

**Things That Make You Go Hmmm...**

Why is it called a building, instead of a built?
Why does sour cream have an expiration date?
Why is it that writers write but fingers don't fing?
Why is a slim chance the same as a fat chance?
Why are quite a few the same as quite a lot?
Why does sour cream have an expiration date?
Why are quite a few the same as quite a lot?
Why is there no egg in eggplant, no ham in hamburger and neither apple nor pine in pineapple?