Robots and the Call for Natural Semantic Systems

By Paul Cohen

First let me thank you for making us—Carole, Clay, Joshua and me—feel so welcome at ISI. We were able to start work quickly, thanks to your support and advice, and we really appreciate the sense of community here. You probably know that ISI has a Center for Research on Unexpected Events (CRUE) and that Yigal Arens was its founding director. Yigal graciously handed it off to me, and I have spent the last couple of weeks meeting with many of you to discuss directions, projects and specific technologies. This has been enormous fun! In a future issue of the InSiDer, I will summarize the themes of these meetings and discuss the future of CRUE. Here, however, I’d like to describe one of my research interests. It is philosophy, but with practical, even economic, consequences.

A defining characteristic and a great accomplishment of artificial intelligence is its reduction of semantics to syntax. As Haugeland put it, “If you take care of the syntax, the semantics will take care of itself.” An early, celebrated example of syntactic faux-meaning was Weizenbaum’s Eliza system. You could chat with Eliza by typing sentences, to which Eliza would respond with comments that ranged from banal to bizarre to insightful. Eliza did not understand the conversation, except in a functionalist sense. Consider an Eliza-like dialog:

Me: "You never take out the trash."
Eliza: "Would you like me to take out the trash?"
Me: "Yes!"
Eliza: "Ok."

Needless to say, Eliza won’t start to take out the trash and wouldn’t, even if it were connected to a robot, because it has no clue what “take out the trash” means. All it knows is that the syntactic (Continued on page 8)

From Here to Eternity: Crossing the Country in a Minivan

By Carole Beal

When friends and colleagues ask why Paul Cohen and I decided to leave our comfortable faculty positions at a shabby but respectable state university back East to come to ISI, the answer is easy: The wonderful research opportunities! The great new colleagues! The fabulous location! What is less easy to explain is why we got from Amherst, Massachusetts, to LA the way that we did.

Then Paul announced one morning over coffee that his lifelong dream was to drive across the country with his family. Okay, I was with him so far; the two of us actually did the trip to Amherst 20 years ago after graduate school, dead broke and driving a Pinto, and I figured it couldn’t be that much worse in a minivan with a couple of credit cards and a kid. But then he added that he wanted to camp (as in, tents, sleeping bags, and dank and stinky restroom facilities half-a-mile away in the bug-filled dark).

I recall staring at him over my coffee cup thinking that this was one of those defining moments in a marriage when all you could manage to do was not shriek, “Are you absolutely insane?”, a temptation that became even stronger when it transpired that we would be bringing not only our own seven-year (Continued on page 6)
## October 2003

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 Manisha’s Birthday!</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

### November 2003

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### SCHEDULE OF EVENTS:

#### October:
- October 1: Manisha Joshi’s Birthday!
- October 1: IUI Submissions Due!
- October 6: Hans Chalupsky’s Birthday!
- October 10: Rahul Bhagat’s Birthday!
- October 14: Lei Qu’s Birthday!
- October 14: Snehal Thakkar’s Birthday!
- October 18: Lei Ding’s Birthday!
- October 22: Alma Nava’s Birthday!
- October 23: Jonathan Grael’s Birthday!
- October 23: Herwin van Welbergen’s Birthday!
- October 28: ICAPS Submissions Due!

#### November:
- November 2: Franz Och’s Birthday!
- November 4: Ning Wang’s Birthday!
- November 6-7: ISD Retreat!
- November 7: ICRA Submissions Due!
- November 12: Erin Shaw’s Birthday!
- November 12: Radu Soricut’s Birthday!
- November 13: Patrick Pantel’s Birthday!
- November 14: Andrew Philpot’s Birthday!
- November 18: Tom Murray’s Birthday!
- November 23: Bryant Huang’s Birthday!
- November 26: Ulf Hermjakob’s Birthday!
- November 26: KR Submissions Due!
- November 27-28: ISI Closed!
- November 29: Richard Whitney’s Birthday!
**SCHEDULE OF EVENTS:**

December:

- December 8: Shashi Kolar’s Birthday!
- December 8: Andre Valente’s Birthday!
- December 15: Shou-de Lin’s Birthday!
- December 15: Varun Ratnakar’s Birthday!
- December 20: Stefan Decker’s Birthday!
- December 23: Kasra Mogharei’s Birthday!
- December 25: Wei-Min Shen’s Birthday!
- December 25-January 6: ISI Closed!
- December 28: Dragos Munteanu’s Birthday!

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Shashi’s Birthday! Andre’s Birthday!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Shou-de’s Birthday! Varun’s Birthday!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kasa’s Birthday!</td>
<td></td>
<td>Wei-Min’s Birthday! ISI Closed</td>
<td>ISI CLOSED</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISI CLOSED</td>
<td>ISI CLOSED</td>
<td>ISI CLOSED</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ISD Tidbits**

- Congratulations to Philipp Koehn, who passed his Ph.D. Defense in July!
- Congratulations to Ion and Maria Muslea, who recently became parents! Hana Malena Muslea was born on August 4 and weighed in at 20 inches and 6 lbs 14 oz! Parents and child are doing well.
- Lauri Grier has left the building! Lauri has moved with her family to the upstate New York area, and Lauri is looking forward to the cold, harsh winter! Janice Cole has joined our group and will be taking over Lauri’s position. Lauri will be greatly missed here.
- Congratulations to Yolanda Gil, who was recently elected Councilor of the American Association of Artificial Intelligence (AAAI)! Each year AAAI members elect four new councilors to serve three-year terms. Craig Knoblock is also a current member of the Council. The AAAI Executive Council is responsible for the formulation and maintenance of the policy of the association. If you have any ideas or suggestions for the role of AAAI, Yolanda and Craig would be very interested to hear from you.
- Uli Germann has left ISI to pursue his Ph.D. at the Computer Science Department of the University of Toronto. Uli can still be reached at his ISI email address for a bit longer, and then you can get in touch with him at germann@cs.toronto.edu.
- Congratulations to Lewis Johnson, who was re-elected to the Executive Committee of the International Artificial Intelligence in Education (AIED) Society!
- Anton Leuski will soon be leaving ISI to return to UMass. You can continue to reach him at his ISI email address for a while.
- Congratulations to Shou-de Lin and Hans Chalupsky, whose paper titled "Using Unsupervised Link Discovery Methods to Find Interesting Facts and Connections in a Bibliography Dataset" was selected 2nd place for the open task in ACM KDD Cup 2003 (http://www.cs.cornell.edu/projects/kddcup/results.html)!
- Congratulations to Shou-de Lin and Craig Knoblock, who received the best paper award of the 2003 IEEE/WIC International Conference on Web Intelligence, for their paper, titled, "Exploiting a Search Engine to Develop More Flexible Web Agents."
- Mike Fleishman has left ISI to pursue his Ph.D. at MIT. You can reach him there at mbf@mit.edu.
CAROLE BEAL
CBEAL@ISI.EDU
I am a native Californian, who has returned after 20 years in New England. I’m on leave from the University of Massachusetts-Amherst, where I am a professor in the Psychology Department. My primary research interests are in intelligent tutoring systems for mathematics learning in middle and high school students, but I also have worked in other areas of developmental psychology, including children’s eyewitness memory, metacognition and writing development, and the relation of drawing skill and children’s motor development. Along with my spouse, Paul Cohen, another New Face, I have a seven-year old daughter, Allegra. At ISI, I am working with Lewis Johnson’s CARTE group on the Rapid Language and Social Intelligence projects, as well as developing initiatives in K-12 educational technology.

PAUL COHEN
COHEN@ISI.EDU
Paul Cohen was a professor of Computer Science at the University of Massachusetts, Amherst, for 20 years, before joining ISD. Paul serves as Deputy Director of ISD and as Director of the Center for Research on Unexpected Events. His Ph.D. is from Stanford University in Computer Science and Psychology (1983); his M.S. and B.A. degrees in Psychology are from UCLA and UC San Diego, respectively. He served as a Councillor of the American Association for Artificial Intelligence, 1991–1994, and was elected in 1993 as a Fellow of the AAAI. Paul’s projects include Capture the Flag, a wargaming environment; the Robot Baby project, in which a robot learns representations and their meanings sufficient for natural language and planning (see the article in this issue of the InSiDer); and he spends far too much time fiddling with time series algorithms for segmenting sequences, finding predictive patterns, and the like. With Clayton Morrison, Erin Cannon and Carole Beal, Paul works on the “Maps for Verbs” project. Children and adults are shown little movies of interactions between blobs on a surface, they provide verbal descriptions, and we find that verbs and other words related to action cluster nicely, according to the dynamics of the movies. Paul is developing the Bayesian Blackboard architecture for intelligence analysis with Clayton and Joshua Moody, and testing it in the wargaming environment with his lab at the University of Massachusetts. There will soon be a web-based version of the Hats terrorism simulator, developed by Paul, Clayton and Joshua, and a standalone version is available now for all you homeland security, social network, intelligence analysis people. Erin, Carole and Paul are also working on a longitudinal study for DARPA of how office assistants learn; it has yielded a theory of “procedural learning” in which domain knowledge is acquired and organized around procedures. Paul also lectures and writes on empirical methodology. He is writing a report on evaluation for the IPTO office at DARPA, and he wrote Empirical Methods for Artificial Intelligence, a textbook published by MIT Press in 1995. He’s happy to discuss methodological issues with anyone at ISI. Other books include Volumes III and IV of The Handbook of Artificial Intelligence with Edward A. Feigenbaum and Avron B. Barr; and Advances in Intelligent Data Analysis with Xiaohui Liu. Paul is working on a book about Robot Baby.

JANICE COLE
JCOLE@ISI.EDU
I was born in Detroit Michigan sometime during the disco era. During its halcyon days, Detroit was known for Motown and the Big Three car companies. Despite its current image and the fact that, geographically, the state resembles a mitten, Michigan was a decent place to grow up.

Decency and mittens aside, I left Michigan right around the time that the internet was being born, to attend the University of Southern California. I studied Business Administration and found that the part of the curriculum that I was most drawn to was finance.

After graduation, I worked in retail...
banking. I moved on to work at an investment bank for six years. During my free time, I pursued and captured a Masters of Business Administration from Pepperdine University.

One of my favorite pastimes is travel. Over the past few years, I have been fortunate enough to visit Great Britain, France and Italy. As time permits, I am looking forward to visiting other countries as well as traveling within the United States.

Now I find myself back at USC, and I couldn’t be more content. My experience with USC has been filled with great people, great courses and brilliant memories, and I’m sure that my new role within ISI will continue that tradition.

STEVEN DENEEEF
SDENEEEFE@ISI.EDU

I am a first year Ph.D. student, and I am still trying to figure out what I am doing. I am interested in Natural Language Processing (especially Machine Translation), and my advisor is Kevin Knight. I don’t have a lot of formal experience in NLP, MT or even AI, but I am thoroughly enjoying learning all about these things, and I am looking forward to making my own contributions soon!

Before starting this program, I was a professional software developer for seven years. I worked for a software company called ESRI, creating Geographic Information Systems software. The majority of my work was in Address Geocoding, which has similarities to Natural Language Processing, but is much simpler.

I am also recently married (since Feb 19, 2002) and have no kids (yet). My wife, Melissa, and I like to watch movies, spend time outdoors and are very involved at our church (especially in the area of music). I am learning to play the guitar and like to sing, too. And I like electronic gadgets (PDAs, Laptops, etc.). I don’t know what else to say… you’ll have to stop by and visit me sometime to get to know me more.

JACOB EVERIST
EVERIST@ISI.EDU

I completed my B.S. in Computer Engineering at Oregon State University in 2002 and then stayed at OSU for a year of post-baccalaureate studies and soul-searching. I am currently resolute in my pursuit of graduate studies in robotics, with my ultimate goal of a Ph.D. in computer science. I am currently enrolled in the M.S. program at USC.

I am working in Wei-Min Shen’s Polymorphic Robotics research lab on the SOLAR project. SOLAR is a project to develop autonomous robots to assemble space structures in-situ.

My interests include robotics modeling, self-reconfigurable robotics, self-assembly, distributed robotics and human-robot interaction.

BRYANT HUANG
BHUANG@ISI.EDU

I love language with a passion. In fact, it has been the passion most of my life to learn languages and learn how languages work. I have also had a strong interest in computers from a very young age. I would consider programming fun, while my friends would immediately run! My childhood fascination in those two areas has now become my professional field of interest.

I am so excited to be here at ISI as a new graduate research assistant. I have the wonderful opportunity of working with Kevin Knight in my favorite area, machine translation.

This past June, I graduated from UCLA with a Bachelor’s degree in (as you’d expect) Linguistics and Computer Science. Currently, I am a first-year Master’s student in Computational Linguistics at USC. I hope to be able to contribute my linguistics knowledge where it is needed.

In the little free time that I have, there are so many things that I love to do. Music: guitar, singing, songwriting. Sports: basketball, tennis, snowboarding, baseball. Learning languages, reading books. Relaxing with friends, mentoring younger boys. Rollerblading, bike
old daughter, Allegra, but Paul’s half-sister Zana, also seven-years old—they share the same birthday (it’s a long story)—and that he would be out of town for the last two weeks before the departure, leaving me to coordinate the entire move and the camping trip preparations (not that I’m bitter or anything). Actually, to be fair, Paul and Allegra planned the entire two-week trip together, doing much research on interesting places to visit and the best routes, and formulating precise time-distance-mileage estimates of daily progress.

When the departure day arrived, Allegra and I left the movers to finish packing the truck and drove to the Hartford, Connecticut, airport, where we met Paul and Zana, who greeted us by announcing that she had “barfed all over the plane,” due to motion sickness. This did not bode well for a 3000-mile trip in the back seat of a minivan, but we tucked a handful of airsickness bags in the glove compartment, and we headed west.

As a special treat to get us off to a good start, we stayed the first night at an off-brand Howard Johnson’s outside of Albany, New York (ask me about the motorcycle gang on the stairs), and then it was on to Niagara Falls. We had given the girls the choice of visiting the chocolate factory at Hershey, Pennsylvania, or of visiting Niagara Falls and eating a Hershey’s chocolate bar on the Maid of the Mist boat: They chose the boat, perhaps not quite understanding that they would have also gotten chocolate at Hershey’s. Impressions from Niagara Falls are, basically, that a heck of a lot of water goes over the edge every second or so. One could say more, but that does about cover it.

Our first night of camping went well; we managed to put up the tent without giving away the fact that we had only done it once before (in our living room), and Paul, who is a fabulous cook, somehow managed to prepare pasta with white beans and rosemary on a one-burner propane stove. (This was merely one in a series of fabulous meals, including a Spanish tortilla potato omelet and pork tenderloin with ancho chili glaze).

The next morning, we headed for Ohio, where we had booked tickets to see an outdoor theater performance about Blue Jacket, a teenage boy in the pioneer days, who became a Shawnee chief. I am not sure how historically accurate it was, but there was much galloping of horses (which can be adopted during the off-season, if anyone is interested), shooting of flaming arrows, burning of forts, and flexing of feather-decorated pectoral muscles, not to mention the dramatic sunset and huge thunderclouds behind the outdoor stage, just as Blue Jacket engaged in the final epic struggle with Daniel Boone. Dazzled by the pectoral muscles and completely missing the significance of the dark, looming clouds, we got back to the campground to find about three inches of water in the tent. Resisting the urge to find a Holiday Inn Express, we mopped it up and slept in damp sleeping bags. (I want points for that.)

It became clear over the next day or two that thunderstorms were going to be a big problem, so in desperation to get out of the mid-west, we started driving harder, longer and faster than planned and gained two days on the original itinerary, which we spent at the Arch in Saint Louis (where they sell hyper-claustrophobic people like me tickets to get to the top, without mentioning, until it is far too late, that you have to sit with three other people for six minutes in an air-conditioned elevator car the size of a household refrigerator, and then do it all over again to get down, all the while wishing that you’d thought to bring one of those airsickness bags from the car); touring caverns in Missouri (representative comments from some younger members of the party included, “boring” and “a lot of standing around in the dark waiting for someone to shine a flashlight on a rock”); and visiting the Steamship Arabia Museum in Kansas City, which presents the entire cargo of a fully loaded 1850’s steamship that sank in the Missouri River and was later excavated out of a cornfield (the river changed course) by five guys who wanted to get rich, who sunk all their money into the project, and who then decided that the history was more important and that they would keep the collection together (note: they say that there are at least 200 other such steamships under the corn just waiting to be dug up, some of which are rumored to have been loaded with barrels of Kentucky bourbon, which would now be, oh, let’s say, 150 years old...and quite valuable).

Anyway, after a side detour to Dodge City, Kansas, where the girls were enthralled by a mock gunfight in the street and then by a saloon song and dance routine, which included Miss Kitty (who was either very brave or very stupid, sitting in Paul’s lap and singing, while toussling his hair), we stopped in Denver to visit my sister and her three small children. We observed that, yes, although statistically unlikely, it is quite possible for four out of five kids to simultaneously decide to pick their noses when asked to smile nicely for a family photo. We then spent several days driving through Utah, admiring the vast...
Spaces and colorful rock formations. Spirits were buoyed by several odd coincidences, including a double rainbow at Arches National Monument (two little girls born on the same day and two rainbows—get it?) and my wistful murmur, as we drove into Moab, that it would be awfully nice to stay at a Holiday Inn Express that night and to have a hot shower, without having to insert quarters and step around dead spiders on the concrete floor, exactly 15 seconds before we drove past a large “Grand Opening! Holiday Inn Express-Moab!” sign.

Later, we had a wonderful visit to the Anasazi State Park Museum, which just happened to be holding workshops on the lawn that day about historical crafts: The girls learned to make beads from soapstone and to string them on to cord made from dogbane, while Paul engaged in deep debate with a mountain man guy—who made his own hunting bows—about whether ash, birch or alder is most flexible, and while I kept a close eye out for Mulder and Scully. (I’m still wondering about that whole Melissa thing.)

We decided to skip the Grand Canyon on the grounds that the girls would get there, look down, think, “big hole” and be totally unimpressed, and that I would go out of my mind with anxiety thinking that they would fall over the edge to a gruesome death every single second that we were there. Also, the mule trips down into the canyon did not take children that young (mules are not fools). Instead, the girls and I took a sedate horseback-riding trip along the rim of Bryce Canyon. The trail guide muttered, “Off like a herd of turtles,” as we set off, strapped securely to old, slow, very tired horses. We also stopped in Zion National Park, which was beautiful but so hot (108 degrees) that the girls did nothing but drink water nonstop; fortunately by this point, they were experts at getting up and peeing by the side of the tent in the middle of the night and then looking for the damp spot as we took the tent down the next morning. They were pretty starved for entertainment by that point in the trip.

Our final stop was Las Vegas, which I had thought would be the perfect glitzy ending after two weeks of camping. We found an incredibly cheap internet rate for a room at the Excalibur, a hotel characterized by an unfortunate decorative interpretation of the King Arthur legend, and then we blew the savings on front row seats to the Tournament of Kings, a dinner theater extravaganza in which you sit around what is basically a large indoor corral and watch knights on horses battle it out with each other, while you eat a historically accurate meal of chicken, potatoes and dragon’s blood (Pepsi). The historical accuracy part was that you have to eat with your hands—no utensils provided—and, I guess, the dust on the food from the horses’ hooves, as they galloped back and forth about ten feet away from your plate.

On the final morning, as we prepared to leave for San Diego to drop Zana off with her family, we decided to splurge on a breakfast of Krispy Kreme doughnuts, which we had heard about, but, being from the East Coast, had never tried. We noticed that Allegra was looking a bit pale and were momentarily surprised, in a vague sort of way, that she, Mistress of the Two Food Groups (Sweet, Other), only ate part of one delicious, sugar-crusted doughnut. But in the rush to pack everything up, we thought no more about it until she, well, threw up all over the room. Let us just end by saying, remember those airsickness bags in the glove box? They certainly came in handy, as we came over the Grapevine and headed to the coast. So, two weeks after we started, we arrived in SoCal with the only casualties being two pairs of children’s swim goggles and two yellow plastic ducks left behind in Effingham, Illinois. (I know exactly where they are, so if anyone is heading that way, let me know.)

Anyway, was the trip worth it? Um, let me get back to you on that... Well, we definitely have two weeks’ worth of wonderful family memories. For example, when asked what she remembered about the trip, Allegra thought carefully and then replied, “The Hershey chocolate bar on the boat.” I do think the girls learned that a) this country is very large, and b) there is a lot of empty space in the middle. Would we do it again? Oh, maybe in another 20 years... Start saving those little white bags from the plane for us!  

Crossing the Country in a Minivan, continued
form, “You never X” can be responded to with, “Would you like me to X?”.

Fast forward thirty years, and we find that programs are much more sophisticated than Eliza but take similar approaches to meaning. Many problems can be solved by attending to very narrow aspects of meaning or to none at all (medical expert systems, semantic web applications and information retrieval systems come to mind). Problems framed as finding the most likely output given particular input often can be solved by statistical algorithms that do not attend to the meanings of the input and output. The tremendous successes of natural language research at ISI and elsewhere attest to the power of statistical algorithms. So what’s the problem? Why fuss about meaning?

One reason to fuss about meaning is that I don’t think the problem has gone away. We have! It is very expensive to write ontologies and axioms that say what terms mean in anything like the depth and detail that we expect of humans. After decades of fighting this battle in AI, we have become adept at selecting problems that our programs can solve without attending to meaning; or rather, problems for which syntactic operations are good proxies for semantic operations.

Another reason to fuss is what we might call semantic illusion: We provide programs with inputs that are meaningful to us, the programs produce outputs that are meaningful to us, and presto, we think the machines understand the inputs. We have the illusion that information systems understand our queries, because these systems produce results that are semantically related to the queries. But these systems don’t understand, any more than medical expert systems understand terms like meningitis. Expert systems, and many other successful AI systems, are essentially scratchpads for us, the users. We assign meanings to the symbols that these systems shift around. Said differently, the meanings of symbols have no effect on how these systems process them.

I want to contrast conventional functional systems with natural semantic systems. The sense of “conventional” that I intend is “by convention, not idiosyncratic or personal,” and “functional” is used in the sense of functionalism, the doctrine that meanings of symbols are roughly the effects of the symbols on behavior or function. A natural semantic system is one that acquires and maintains meanings for itself. Natural semantic systems clearly are disjoint from conventional functional systems. Humans are natural semantic systems. We learn meanings of mental states and representations, refining these meanings throughout our lives. Natural semantic systems are not mere scratchpads for meanings assigned by exogenous agents (i.e., programmers), nor must they be told what things mean. You don’t require anyone to tell you the meaning of stubbing your toe or finding $100 in your pocket. I don’t have to tell you what this sentence means or which inferences follow from it. You know what events and sentences mean because you have learned (or have learned to work out) these meanings. Because you are a natural semantic system, I expect you to understand this paragraph and quite literally to draw your own conclusions about it. If your conclusions differ from my own, it is not because you have a bug in your program that I need to fix; it is because you maintain a semantic system that overlaps largely, but not entirely, with mine.

Natural semantic systems are rare in AI. A case might be made that reinforcement learning produces natural semantic systems, as the meanings of states (i.e., their values) are learned from a reinforcement signal. Sadly, much work in reinforcement learning tries to coerce systems to learn value functions (i.e., meanings of states) that we want them to learn. This is accomplished by fiddling with the reinforcement signal, the state space and a batch of other factors, until the system produces what we consider to be the correct outputs for a set of inputs. It is ironic that reinforcement learning is capable of producing at least rudimentary natural semantic systems, but that we use it to produce conventional systems.

Conventional functionalism is a good way to build systems with a little semantics that do what we want them to, so most AI systems will, quite appropriately, continue to be based in conventional functionalism. The issue is not whether we should stop building such systems, but whether we can actually build conventional functional systems with more than a little semantics.

Let us also recognize that, for a given specific task, the semantic content acquired naturally by an intelligent agent can be duplicated or mimicked by a sufficiently motivated knowledge engineer; so the issue is not whether natural semantic systems can think or mean things that conventional functional systems cannot think or mean, but whether there are enough knowledge engineers on the planet to build a conventional functional system that understands as much as we do.

The crux of the argument against conventional functional semantics and for natural semantics is this: Conventional functional systems are entirely syntactic, and all meanings are exogenous, so to build one, you have to design a system whose syntactic operations produce results that you can interpret as meaningful. This design task is very expensive; so expensive, in fact, that we have not been able to build conventional functional systems for semantically-deep tasks. Programming a computer is a paradigmatic example of this design problem. If you want meaningful outputs, you have to write programs whose syntactic operations are meaningful to you. If this were easy, you’d never have to debug your code. Doug Lenat and John Seely Brown recognized this problem in a paper called “Why AM and Eurisko Appear to Work” (AAAI, 1983). The AM system discovered many concepts in number theory, but when the AM approach was tried in other domains, it didn’t work well. Lenat and Brown concluded that AM worked because syntactic Lisp operations on Lisp representations of mathematical concepts often produced meaningful new mathematical concepts, whereas syntactic Lisp operations on Lisp representations of other kinds of concepts rarely produced meaningful new concepts:

“It was only because of the intimate relationship between Lisp and Mathematics that the mutation operators ... turned out to yield a high "hit rate" of viable, useful new math concepts. ... Of course we can never directly mu-
tate the meaning of a concept, we can only mutate the structural form of the concept as embedded in some representation scheme. Thus there is never a guarantee that we aren’t just mutating some ‘implementation detail’ that is a consequence of the representation, rather than some genuine part of the concept’s intentionality.” (Lenat and Brown, AAAI, 1983, p.237)

So, Haugeland is correct to say that good old-fashioned AI tries to “take care of the syntax, and the semantics will take care of itself,” but taking care of the syntax is very hard!

Let’s be clear about why conventional functional systems are hard to build. Lenat and Brown are correct that all computer-based systems, including natural semantic systems, operate on the “structural form” of concepts, not directly on the meanings of concepts. The problem is maintaining a correspondence between syntactic forms and their meanings, and the system itself checks whether syntactic operations produce meaningful results. Suppose a robot is planning to move to an object. Syntactically, it generates a plan to rotate the left wheel in one direction and the right wheel in the other direction. When it executes this plan, it spins around. The meaning of this plan is, among other things, that it does not achieve the robot’s goal of moving toward the object. But we do not have to tell the robot this, because the robot maintains the semantics of its plans, itself.

In natural semantic systems, operations on data structures are also syntactic, but the system itself is responsible for maintaining the correspondence between syntactic forms and their meanings, and the system itself checks whether syntactic operations produce meaningful results. Suppose a robot is planning to move to an object. Syntactically, it generates a plan to rotate the left wheel in one direction and the right wheel in the other direction. When it executes this plan, it spins around. The meaning of this plan is, among other things, that it does not achieve the robot’s goal of moving toward the object. But we do not have to tell the robot this, because the robot maintains the semantics of its plans, itself.

This checking, and the subsequent correcting, is very expensive. I call this the semantic babysitting problem. Babysitters everywhere know that, sooner or later, the kid is going to do something stupid. Vigilance is the only way to prevent injury. Programmers understand that, sooner or later, the system is going to do something they didn’t intend (not a syntactic error—compilers can catch those—but a semantic one), and vigilance is the only way to catch it.

I think robots will be among the first machines to learn the meanings of representations and to not require semantic babysitting. This is why Carole Beal, Clay Morrison, Tim Oates and I collaborate on The Robot Baby Project and why you’ll be seeing robots roaming the ninth floor in the coming months. Robot Baby is truly in its infancy, but has already made some progress toward learning word meanings. The next step, I think, is to learn deeper meanings (e.g., not only that “forward” means positive rotational velocity, but also that forward and backward are antonyms, and that moving “forward” is one way to get to a goal). Further along, we need to eliminate the AI programmer from Robot Baby’s development. It has always bothered me that Artificial Intelligence requires very intelligent researchers to design automata that don’t have the slightest clue what they are doing or why. I think we need a theory that explains what good AI programmers are doing when they design representations. If I understood how to design representations—how to make formal objects and operations stand for things in the world and operations on them—then perhaps I could have a robot design its own. And this would make the robot autonomous in a way that few are, today.
Research Contracts in ISD

Congratulations to Ed Hovy, who received an NSF contract for his project titled, “Collaborative Research: Interlingual Annotation of Multilingual Text Corpora” (IL-Annot).

Designing an Interlingua—a neutral representation of text meaning that sits between languages and can be used to facilitate machine translation and other multi-lingual applications—has been a dream of computational linguists for several decades. Despite progress on several fronts, a number of key phenomena stubbornly resist standardization. They include: case roles (can one list the most basic roles associated with each action, and each object?), aspect (for example, what does the continuous form in English mean?), discourse connectives (what is the general principle underlying “but” and “however,” for example?), etc.

This project is a collaboration with people from six universities around the country. Each one will propose annotations to describe these phenomena, using his or her language of expertise, and meet periodically to try to come up with a representation that supports at least these six languages adequately.

Congratulations to Paul Cohen and Clayton Morrison, who received an NIMA award for their project titled, “A Collaborative Laboratory Environment for Studying Shared Knowledge Creation” (COLAB).

An architecture for interpreting intelligence data (AIID) will serve as a laboratory in which to study how human analysts work together. AIID is based on the Bayesian Blackboard, developed by Cohen and his colleagues at the University of Massachusetts. The COLAB contract is from the DCIA Postdoctoral Fellowship program. Clayton Morrison wrote the proposal and is the recipient of the award. Morrison works with Cohen in the Center for Research on Unexpected Events (CRUE).

Congratulations to Ed Hovy, who received an NSF contract for his project titled, “Automating the Integration of EPA Databases” (EPA-Air).

As with many large organizations, the Government’s data is split in many different ways and is collected at different times by different people. The resulting massive data heterogeneity means that government staff cannot effectively locate, share or compare data across sources, let alone achieve computational data interoperability. A case in point is the Air Quality Management Office of the California EPA, which every year has to integrate the air quality readings from its 36 regional offices and send them to US-EPA in North Carolina (which in turn has to integrate the data from all 50 states and from neighboring countries).

This project will employ the new statistical algorithms developed since 1990 in the Machine Translation community to discover correspondences across comparable datasets at all levels. If our automatically learned mappings are effective, we should be able to significantly reduce the amount of manual labor required in database wrapping.

Congratulations to Jose-Luis Ambite, who received an NSF contract for his project titled, “Dynamic Composition of Web Services for Goods Movement Analysis and Planning” (ARGOS).

The goal of ARGOS is to develop a flexible data query and analysis system, based on the web services paradigm. As an application domain, we will examine several goods movement planning problems and their effects on spatial urban structure.

Many scientific problems can be modeled as a workflow that includes information gathering and processing operations. We propose a unifying framework where these operations are modeled as web services and the scientific workflows as compositions of web services. Argos provides graphical tools for manual specification and composition of web services, as well as automatic composition based on expressive web service descriptions for given application domains (such as transportation planning).

We will use Argos in a metropolitan transportation planning scenario in consultation with our advisory team, that includes government practitioners from the California Department of Transportation, the Los Angeles County Metropolitan Transportation Authority, the San Bernardino Associated Governments, the Southern California Association of Governments, the Port of Long Beach, and others. The scenario analysis will allow evaluation of Argos in terms of its expressiveness, utility as a transportation planning tool, and ease of use by developers and practitioners.

Congratulations to Paul Cohen, who received a DARPA and Air Force Research Lab (AFRL) contract for his ORG-EVAL project.

ORG-EVAL is two projects rolled into one, the first a study of how human office assistants learn procedures and domain knowledge, the other to develop evaluation methods for IPTO projects.

New PowerLoom Software Release

By Tom Russ

The Loom group is pleased to announce that we have just released a beta version of our PowerLoom™ 3.0 knowledge representation system. This release marks a major milestone in our efforts to provide a follow-on system to the venerable Loom™ knowledge representation language. Among the major new elements in this release are the first draft of the PowerLoom manual and a new programming interface for using the system. The 3.0 release also represents a major improvement in performance and capabilities over the previous releases.

PowerLoom provides a language and environment for constructing intelligent applications. PowerLoom uses a fully expressive, logic-based representation language (a variant of KIF), and it uses a natural-deduction-style backward and forward chainer as its inference engine. The inference engine is not a complete first-order theorem prover, but it can handle complex rules, negation, equality reasoning, subsumption, and restricted forms of higher order reasoning. PowerLoom has a classifier that is able to classify descriptions expressed in full first order predicate calculus. PowerLoom uses modules as a structuring device for knowledge bases, and lightweight worlds for classification and hypothetical reasoning.

The latest release is available in Common-Lisp, C++ and Java versions. More details can be found at http://www.isi.edu/isd/LOOM/PowerLoom.

Loom and PowerLoom are trademarks of the University of Southern California. We are currently in the process of registering these trademarks.
In the world of K-12 education, “accountability” is the current hot buzz-word: Parents, politicians and educators want evidence that the classroom instruction being experienced by children is effective. In many cases, for better or for worse, this means frequent high stakes testing, even for grade-schoolers. We could argue indefinitely about whether test scores are an appropriate outcome measure, but given the current reality of high stakes testing, it is critical to find ways to help all students prepare to do as well as they can. Intelligent tutoring systems, with the ability to individualize instruction for students with a range of needs and characteristics, offer a promising solution.

In the Wayang Outpost project, currently supported by an NSF grant awarded to me and Beverly Woolf (my colleague at UMass-Amherst), our goal is to create an intelligent tutor system to help high school students improve their scores on the SAT-Math exam. We focused on the SAT-Math exam because this test has long been known to yield large gender and ethnic group gaps in performance, which, in many cases, do not correlate with students’ classroom performances, and because the SAT serves as a “critical filter” for college admissions, scholarships and entrance into STEM (Science, Technology, Engineering and Mathematics) programs.

You might wonder, aren’t there plenty of commercial test preparation programs out there already? Yes, but when we reviewed a number of such programs, we discovered that they focus primarily on providing information about the structure of the SAT (e.g., how many sections) and about test taking strategies (e.g., move on if you can’t solve the problem in 30 seconds); they don’t actually teach you how to do the math! Our hypothesis was that the real problem for many students, particularly young women, was that, when they were confronted with unfamiliar, often challenging math problems on the test, they were not transferring the math skills that they learned in the classroom effectively. As it turned out, when we analyzed a large set of SAT problems provided to us by the College Board, we discovered that in many, many (read: almost all!) cases, the problems can be solved or, at least, made much easier than first appears, by mental rotation, visualization, estimation, or other spatially-oriented strategies: all approaches that females tend to be less likely to use than males.

Armed with this discovery, we created Wayang Outpost, a web-based intelligent tutoring system for high school students. The tutor integrates mathematics with an environmental biology theme, to help students see the importance of math to science, and introduces the student learner to a virtual village, based on an orangutan research station in Borneo. The teaching decisions and learner model run on a server at UMass and interact via XML, with Flash movie files that contain the problems and characters. The Flash files are downloaded to the user’s local machine in real time; we use the learner model to anticipate what help the user will request and download it before the user asks for help.

Characters in the Outpost are based on Indonesian shadow puppets (the art of Wayang: a virtual world designed to convey important lessons to young people, hence the tutor’s name) and also on real female scientists and researchers who serve as role models for young women, including Ann Russon at York University; Beth Fox at Princeton; Laura Debnar at the San Diego Zoo; and Lori Perkins, Zoo Atlanta. The tutor includes on-line assessments of the student’s cognitive abilities (math fact proficiency; spatial cognition), as well as pre- and post-tests of SAT-Math ability. The heart of the tutor is the “SAT Hut,” in which the student can work with SAT problems and receive help that either shows an algorithmic solution (what female students tend to do on their own—effective, but often slow and prone to error) or a visually-oriented solution (what male students tend to do—fast, often effective).

Wayang Outpost also includes animated adventures, led by one of the female scientists, which include multi-step, challenging math problems, based on real environmental issues related to the area. For example, in the “Illegal logging” adventure, the student and Lori Perkins observe an illegal logging operation and create a graph showing the gap between the number of logs that the truckers are permitted to harvest versus their actual harvest over the course of the season (using proportional reasoning); subsequently, a heavy rain causes flooding (because when the trees are gone, the thin soil is easily washed away), and the student must estimate the flooded area and the likely migration of the orangutans that will result. The adventures are designed to provide a measure of transfer: if students have learned anything from the SAT problem tutoring, we hope that they will be able to figure out how to solve challenging problems that might be similar to what a scientist in the field might encounter.

So far, we have conducted several evaluation studies with high school students in two school districts in Western Massachusetts, and we’re encouraged by the results. Students do improve their mock SAT scores after working with the tutor, and one intriguing result was that women improved most of all when provided with visually-oriented tutoring strategies; in fact, they outscored men on the post test. Students also reported liking the help and the adventures, giving us comments like, “I really like the animations. The way I learn is by seeing things working out, so it was very good.”; “I really liked the math adventure.”; “I liked it because it was like a computer game; it made math more fun.” Our first evaluation study to include all the animated adventures is scheduled for later this month, so stay tuned for the results in a future issue!
riding, ice skating, rowing, being out in magnificent nature! Eating (be sure to invite me!). If there’s something I haven’t tried, please introduce me. In addition, I am very active in serving at my church.

CLAY MORRISON  
CLAYTON@ISI.EDU

My name is Clay Morrison, and I am a Postdoctoral Fellow working with Paul Cohen. The move to the Los Angeles area is somewhat of a homecoming for me, although it is entirely new for my family. I was born and grew up in the Altadena and Pasadena area, while my wife, Heather, grew up in upstate New York (where we met), and my daughters Maya (three years) and Audrey (eight months) were born in western Massachusetts. It is nice being back near the beach and "real" mountains.

My academic career is somewhat convoluted, but has followed particular themes—in particular, an interest in how we might build intelligence systems capable of autonomously learning and developing. I attended Occidental College (in Eagle Rock), receiving my B.A. in 1992, majoring in Cognitive Science. During the latter half of my studies, I realized that much of my interests lie in the philosophical foundations of cognitive science. I entered the program in Philosophy, Computers and Cognitive Science in the Philosophy Department at Binghamton University, New York, where I completed my M.A. and Ph.D. in philosophy. My thesis work developed a critique of the role of representation in computational models of analogical cognition. Near the completion of my Ph.D., I had the good fortune to meet Paul Cohen and found that we shared many similar interests. I spent the past four years at the University of Massachusetts, Amherst, where I worked as a lab manager in the Experimental Knowledge System Laboratory and filled in my computer science background working on a computer science M.S.

I am currently working on developing AIID, an Architecture for the Interpretation of Intelligence Data. AIID is a Bayesian blackboard system that constructs an interpretation of an unfolding scenario by incrementally building Bayesian networks through the interaction of blackboard knowledge sources. This year, I will focus on adapting the AIID framework to accommodate human analyst interaction with the blackboard for the study of analyst collaboration; the augmented system is called COLAB.

TOM MURRAY  
TMURRAY@ISI.EDU

I am a new Ph.D. student, working in the NL group with Eduard Hovy. My interests naturally center around language and how to make human-machine interfaces work on human terms, and not the reverse. Question-answering, ontology-building, and multi-modal communication especially intrigue me. I’m currently working on analyzing organizational relationships from e-mail correspondence.

A long time ago, I earned undergraduate degrees in English Literature and Mathematics at the University of Missouri-Columbia, in the state where my hippie parents happened to be when I was born.

After a brief stint teaching English in Moscow (where I regularly was mistaken for a Chechen by gun-toting militia), I decided that writing software at start-up companies was the best way to engender chaos and stress in my life. This path at least allowed me to escape the Midwest to SoCal, and to meet my lovely wife, Vivian Le Tran, a writer and artist.

Fortunately, not all of the dot-coms I worked at were ridiculous; I managed to learn a lot and gained exposure to problems and techniques that finally whetted my appetite to return to school. In June, I quit my job leading the Architecture Group at Evolution Robotics in Pasadena and road-tripped 10,000 miles around the country, completing the metamorphosis to Student.

When I’m not busy doing student things, I like to cook vegan food, explore the world, break and sprain limbs mountain biking, and make noise singing and playing guitar.
New Faces, continued

PATRICK PANTEL  
PANTEL@ISI.EDU

Hello all! I am very excited to join Ed Hovy’s Natural Language group as a Research Scientist. I will be working on text mining and (semi-) automated ontology construction. Recently, I completed my Ph.D. from the University of Alberta, where I worked on mining text for building and extending semantic repositories. You can play with some of my demos on my website: http://www.isi.edu/~pantel.

I have lived in Canada all of my life so, come January, please do not come to me to complain about the weather, if it gets a little chilly ☺. But, please drop by if you need a partner to play tennis, soccer, basketball or whatever sport you like (even hockey!). I also enjoy traveling, hiking and playing piano.

GLADYS SAROYAN  
SAROYAN@ISI.EDU

I am involved in the Tactical Language Project with Lewis Johnson. I was born in Lebanon and emigrated to the United States as a teenager with my family. I have a degree in Journalism from San Francisco State University. I wrote regularly for the San Francisco Examiner’s Opinion Page. I am published in numerous other publications, including Anthologies on the Middle-East, edited by the poet Shihab Naomi Nye; "The Space Between Our Footsteps," and "The Flag of Childhood." I speak several languages. I taught Arabic and French at the Berlitz School of Languages. I am currently a weekly columnist for the Daily Breeze Newspaper, San Pedro edition. My husband is a psychologist. I have three children. I own a dog, a bunch of fish and a sickly looking hamster. Unusual for a Lebanese to be harboring animals in her home, but, when in Rome...

HERWIN VAN WELBERGEN  
HERWIN@ISI.EDU

Hi, I am Herwin van Welbergen, an exchange student from the University of Twente (the Netherlands). I’m working on the Social Intelligence project with Lewis Johnson. My current task is to extend the current Natural Language Generator, and to set it up for connection with a Wizard of Oz User Interface.

In my free-time, I like to visit friends, do some sight-seeing or play chess. I’m also trying to learn surfing and sailing. I look forward to meeting all of you!

HANNES VILHJALMSSON  
HANNES@ISI.EDU

My name is Hannes Högni Vilhjálmsson (that’s "William’s son"), and I joined CARTE as a Research Scientist in August, where I’ve mainly been involved in the Tactical Language project so far. My research focuses on nonverbal cues in face-to-face interaction and how these cues can be autonomously generated in interactive animated characters, based on linguistic and social context.

This summer, I received my Ph.D. in Media Arts and Sciences from the MIT Media Laboratory. My thesis dealt with how to improve online collaboration in shared virtual environments, by using a model of face-to-face conversation to animate the articulated avatars representing each of the participants. My desire to apply this work to online learning and virtual classrooms was one of the factors that brought me to CARTE.

My focus on nonverbal expression and virtual creations is not a surprise, given my life-long interest in theatre, puppetry, dance, martial arts and cinematic special effects. Furthermore, one of my dreams as a child was to become a master magician (and this was before Harry Potter!), so coming to this tower of high sorcery is really a dream come true.

I moved from Boston to Los Angeles with my wife Deepa Iyengar. We found a nice place to live in Playa del Rey, right between the wetlands and the beach. Being from Iceland, this proximity to nature and getting enveloped every morning in chilling fog from the ocean, makes me feel right at home. I’m already looking forward to the heavy snow in the winter.

SHUMIN WU  
SHUMIN@ISI.EDU

Hello, my name is Shumin, and I came from Hangzhou, China, a beautiful city famous for the West Lake that attracts hundreds of thousands of tourists every year. I completed my Master’s studies at the University of Western Ontario and served as an R.A. at the National Research Council, Canada. I dabbled in the agents area during my Master’s research, and I learnt that ISI has a very strong team in agents. Then, I came to USC for my Ph.D. in August of 2003, and I am glad to find so many researchers who share my interests. And I love Los Angeles—it is cool! I am interested in creating intelligent user interfaces that can help people with their overwhelming computer-based tasks. I dream of making agents speak, learn and smile. I am currently working with Prof. Lewis Johnson and CARTE, researching agents for educational purposes. This involves various disciplines such as Distributed Artificial Intelligence (DAI), Human-Computer Interaction (HCI), and many other state-of-the-art technologies. Please feel welcome to email me anytime to discuss my research!

Research has taken most of my time, but, I’ve still kept some hobbies since high school, such as music, comics, reading and writing novels, and Chinese opera.
Microsoft recently announced that it would increase its research budget to $6.8 billion per year (in comparison, DARPA only spends about $2 billion). Using this money, Microsoft could hire over 750 thousand graduate students for summer internships. Apparently, they elect not to do this and, instead, hire only a few hundred. Despite the consensus among graduate students that they are the only ones who actually do any work, this under-hiring doesn’t result in seven-figure summer salaries, because Microsoft Research (MSR) also employs several thousand full-time researchers.

Of these several thousand full-time researchers, I had the opportunity to work with one in particular: Eric Brill. Eric, who is perhaps most famous for the "Brill Tagger," one of the more popular part-of-speech taggers in the NLP community, also happens to be a really nice guy.

Eric is a member of the ostentatious-sounding "Machine Learning and Applied Statistics (MLAS)" group (there is also an NLP group in Redmond, but they tend to do 80’s-style NLP, whereas MLAS does more interesting research), a name which is appropriate primarily in the sense that it looks good on a C.V. Eric was previously faculty at Johns Hopkins, but he moved to Redmond a few years ago and has spent a large part of his time investigating web search phenomena.

As it turns out, our friend Bill Gates has decided that web search is a very important topic. He has apparently also realized that MSN Search is terrible and that Google is quite good (in Microsoft’s defense, MSN Search is not run in-house—it is contracted out to Inktomi, which is owned, but not used, by Yahoo). The result of this decision and realization is that funding for web search research is even more unlimited right now than it has been in the past. Bill’s goal, with the help of the MSR folks, is to beat out Google (presumably without simply buying them).

The problem that I worked on over the summer had to do with quickly and robustly producing reasonable, structured results for under-specified informational queries. (An informational query is one in which a piece of information is sought, compared to a navigational query in which a particular page is sought.) While we achieved good results, due to corporate machinery and patent lawyering, I’m not at liberty to say much more than that at this point. If the results of my research are actually used in their search engine, I will first determine whether or not people actually like the system, and, if people like it, I will take credit; otherwise, I will pin it on my obnoxious officemate.

In general, I found working at MSR to be only mildly different than working at ISI. Like here, there’s a high ex-Carnegie Mellon population. In fact, I had Carnegie Mellon connections with two people, one of whom was the person who originally introduced me to NLP, the other who remembered reading my homework assignments, when she was my grader for an introduction to linguistics course at CMU. Of course there were many more CMU folk, but in general these are not the sorts of people you want to associate with for no good reason. The environment at Microsoft was similar to the environment here: interns shared interior offices, while full-time researchers got windows. They are experiencing the same space crunch that we are, and the researcher offices there are actually small in comparison to those here.

There are differences, too: one thing that stood out for me, personally, as being nice about MSR over USC/ISI is that they actually employ pure theory researchers. I was able to use being an intern as an excuse to bug the machine learning and algorithms people to teach me things that I had no opportunity to learn here.

Of course, it’s not all roses. For one, you’re forced to use Windows, which is initially painful. Due to this, sharing is much less well-developed than on a Unix-based system, and various tools that NLP people expect to have access to (sentence splitters, tokenizers, taggers, parsers, etc.) are simply not available. And since MSR is not a university, it is often not simply a matter of downloading and compiling these programs (if lawyers don’t stop you, Window’s impoverished build environment probably will). In my project, I ended up not needing any such tools, but I know of at least one other intern from ISI who was underwhelmed by these conditions.

Overall, I enjoyed my time in the pacific northwest. Seattle summers are nice, if at the expense of bad weather the rest of the year. The people there were friendly and quite intelligent. There were little annoyances (like Windows and the bureaucracy that forced us to write “goal” sheets and to go through mid- and end-of-summer reviews), but overall it was an enjoyable time. They treat their interns well, even if they don’t give us seven-figure salaries.
Thought for Food

By Deepak Ravichandran

With many new grad students arriving at ISD at the beginning of this academic year, the most common questions asked by students during the lunch hour are: "Where shall we go to eat?" and "What is a good and cheap place to eat?". The focus of this article is primarily to help new students in answering such questions.

Since time immemorial, it is a well-known fact that graduate students are poor. In fact, some recent studies have shown that grad students live perilously close to the national poverty line. Hence, the obvious choice for a good place to have lunch would be a cheap fast-food place. There are at least 20 fast food places close to ISI. Which one to choose?

There are several criteria that one can consider, when choosing a fast-food restaurant. An AI geek would think that the objective function (criteria) would be to maximize calories per dollar (cal/dollar). In this case, McDonald’s becomes the numero uno choice. (A Big Mac combo easily contains 1200 calories for $3.50!)

But, some studies have shown that the average grad student of the 21st Century is way more health conscious than those of the 20th Century. Then, a reasonable criterion for selecting a restaurant would be to minimize calories per dollar (cal/dollar). Subway easily wins the competition here. (Go to http://www.subway.com for nutrition details).

A new survey conducted at ISD revealed that grad students would prefer to not eat at Subway 86.9% of the time. "Subway sucks," says an AI-grad alumnus, Jay Modi, in an email interview from Pittsburgh, PA.

After careful thought went into several such considerations, I came up with the following criteria for fast-food restaurant selection (not in any particular order):

1. Food should taste good.
2. It should be cheap.
3. It should contain an optimum amount of calories.
4. It should be close to ISI.

The top 5 fast food restaurants that made the final cut are:

**Uncle Darrow’s**

Very few people are aware of this great place at the intersection of Washington & Lincoln. It has awesome tasting Cajun food, loaded with calories. My particular favorite is the Jambalaya and the Zeek wrap.

Taste ***

Cheapness ** ($7-$9)

Healthiness *

Proximity *

**Chipotle**

This Mexican fast-food place right across the street from ISI is a good place for grabbing a quick lunch. The selection is quite poor but they have good rice and beans.

Taste **

Cheapness **** ($5-$6)

Healthiness ***

Proximity ****

**Souplantation**

Souplantation is an awesome buffet-style place for soup and salad. It is also a haven for vegetarians. They have a good variety of salads, soup and desserts to choose from.

Taste ****

Cheapness ** (free!)

Healthiness ***

Proximity ****

**Free Food**

Nothing beats free-food, though. AI-Grad Philipp Koehn, from his beach-facing apartment in Venice, CA, tersely puts it: "Free food is the staple food of all grad students." There are several opportunities for free food: free ISD lunches, free AI-Grad lunches, left over food from project meetings, etc. You only have to keep your eyes open for such events.

Taste ***

Cheapness ***** (free!)

Healthiness ***

Proximity *****

The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM KDD Cup 2003).


By Deepak Ravichandran

Methods to Find Interesting Facts and Connections in a Bibliography Dataset. The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM KDD Cup 2003).


Al–Grads Website!

http://www.isi.edu/isd/ai-grads

E-mail Deepak Ravichandran <ravichan@isi.edu> with any comments or suggestions.

Every grad student should be included, so if you don’t see yourself on the list, email Deepak, so he knows you’re here!