The 2\textsuperscript{nd} International Conference on Knowledge Capture
K-Cap ‘03

Co-chairs: John H. Gennari
Bruce Porter
One person’s overview of

The 2nd International Conference on Knowledge Capture
K-Cap ‘03

Co-chairs: John H. Gennari
Bruce Porter
K-Cap history

- Banff “KAW” workshops: 12 workshops from 1986-1999:
  - Knowledge Acquisition Workshop
  - K-Cap ’03: the 14th KAW?

- K-Cap vs. KAW:
  - Small conference vs. workshop
  - Broader reach – ML and NLP in addition to KA and KB systems
K-Cap history

- Banff “KAW” workshops: 12 workshops from 1986-1999:
  - Knowledge Acquisition Workshop
  - Knowledge Acquisition for Knowledge-Based Systems Workshop
  - K-Cap ’03: the 14\textsuperscript{th} KAW?

- K-Cap vs. KAW:
  - Small conference vs. workshop
  - Broader reach – ML and NLP in addition to KA and KB systems
K-Cap history

- Banff “KAW” workshops: 12 workshops from 1986-1999:
  - Knowledge Acquisition Workshop
  - Knowledge Acquisition for Knowledge-Based Systems Workshop
  - Workshop on Knowledge Acquisition, Modeling and Management
  - K-Cap ’03: the 14th KAW?

- K-Cap vs. KAW:
  - Small conference vs. workshop
  - Broader reach – ML and NLP in addition to KA and KB systems
K-Cap history

- Banff “KAW” workshops: 12 workshops from 1986-1999:
  - Knowledge Acquisition Workshop
  - Knowledge Acquisition for Knowledge-Based Systems Workshop
  - Workshop on Knowledge Acquisition, Modeling and Management
  - K-Cap ’03: the 14th KAW?

- K-Cap vs. KAW:
  - Broader reach – ML and NLP in addition to KA and KB systems
  - Small conference vs. workshop
K-Cap facts & figures

- ACM-sponsored
- Every other year (K-Cap ’05 this October!)
- ~100-150 attendees in ’01 and ’03
- ~25 papers presented in a single track
- Workshops, invited speakers, posters
- Fun, retreat-like locations
K-Cap: The common theme

- An investigation of *Knowledge*:
  - Acquisition
  - Management
  - Representation
  - Reasoning
Knowledge vs. data

A continuum of semantic richness:

- Data (raw numbers)
- Databases (some semantics about the data)
- Concept hierarchies / typologies
- Knowledge bases (both concepts and data)
- Rich ontologies (theories about the data)

If Data and simple DBs are not enough, what is important about knowledge?
Knowledge acquisition

- How do we capture knowledge?
- Knowledge acquisition (KA):
  - Filling in knowledge for expert systems
  - Eliciting knowledge from people
- More formal sorts of knowledge capture:
  - Automatic KA: Machine learning / Data mining
  - KA from text: Text mining
  - Ontology construction methods
Knowledge-based systems

- How do we use knowledge?
  - Reasoning / classification
  - Knowledge management

- How do we store knowledge? (KR)
  - Representation choices (DLs, frames, nets)
  - Ontology design

- How do we share knowledge?
  - Semantic web of ontologies
  - Semantic web reasoners
Well-represented topics/keywords:

- Knowledge Acquisition: 7 papers
- Ontologies: 7 papers
- Knowledge Representation: 6 papers
- Machine Learning: 4 papers
- Natural Language Processing: 4 papers
- Semantic Web: 3 papers

Other topics: Text mining, image annotation, knowledge management…
Best paper session?

- Collecting knowledge from web users:
  - *Collecting commonsense experiences* [Singh & Barry]
  - *Building large knowledge bases by mass collaboration* [Richardson & Domingos]
  - *LEARNER: A system for acquiring commonsense knowledge* [Chklovskii]

- Details? Stay for my next talk…
Panel: *Large Knowledge Capture Projects*

- From Darpa: “Rapid Knowledge Formation” (RKF)
- From Vulcan, Inc: Project Halo, toward a “digital Aristotle”
- From England: “Advanced Knowledge Technologies” (AKT)

What’s happened since Oct ’03?
K-Cap ’03 & ISWC ‘03

- International Semantic Web Conference (ISWC) co-located with K-Cap ‘03 at Sanibel Island, FL
- ISWC: More specific (Web only), but the same knowledge issues:
  - Reasoning
  - Representation
  - Ontology construction & design
Other related conferences

- 14\textsuperscript{th} European Knowledge Engineering & Management conference (EKAW-2004)
- 9\textsuperscript{th} International Conference on the Principles of Knowledge Representation and Reasoning (KRR-2004)
K-Cap and AI

- K-Cap: Not distinct from AI
- K-Cap aims at a particular AI focus: Knowledge
- Thus, K-Cap is a “sister” to many:
  - ISWC
  - EKAW
  - KRR
  - ML
  - AAAI (??)
Thank you!

gennari@u.washington.edu
Two papers from K-Cap ’03:
Capturing Commonsense Knowledge from Web Users

Lead authors:
Push Singh (MIT) and Timothy Chklovski (ISI)

Presenter: John Gennari (UW)
Two K-Cap ’03 papers:

A focus on *commonsense* knowledge:

- Push Singh, Barbara Barry: *Collecting Commonsense Experiences*
- Timothy Chklovski: *LEARNER: A System for Acquiring Commonsense Knowledge by Analogy*

- Research from the MIT media lab
Why common sense?

We may have become too used to putting common sense in that category of “impossible” problems and overlooked opportunities…

[Lieberman, Liu, Singh, Barry (2004)]

- Commonsense Knowledge:
  Lenat and the Cyc project (1985 – 1995 → ??)

- What’s new?
  - Using knowledge in “fail-soft” settings
  - The Web
“Fail-Soft” applications

- Fail-soft environments:
  - Inferences that are only “sometimes” correct are still useful
  - Incorrect inferences are not problematic
- Example: Agents that “push” suggestions to users, augmenting their experiences
- Fail-hard environments:
  - Safety critical
  - Systems that respond to users’ “pull” for information, where accuracy is important
Web users

- Good:
  - Plentiful
  - Cheap
  - “real world”

- Not-so-good:
  - Inaccurate
  - Incomplete
  - Contradictory
  - Malicious (?)
Two K-Cap ’03 papers: #1

- Push Singh, Barbara Barry: *Collecting Commonsense Experiences*
Sept ‘00 – Aug ’03
- 11,000 web contributors
- 600,000 simple assertions of knowledge:
  - “People live in houses”
  - “Running is faster than walking”
  - “Coffee helps you wake up”
  - “Apples are not blue”
  - “The effect of going for a swim is getting wet”
Open Mind Experiences (OMEX)

More than simple assertions and rules:

- Collecting context for knowledge
- Collecting story knowledge (easier for public)
- Collecting cases that enable cased-based learning and inference (Build a large case base)

OMEX capabilities:

- Clone/Create a story (template-based)
- Explain a story
- Judge a story
- Fix/Repair a story
Story templates

- Based on Lernert’s *plot units* e.g.,
  - Competition
  - Failure
  - Motivation
  - Reneged promise
  - Honored request
  - Double-cross
OMEX ‘explain a story’ w/ templates

Please supply at least one fact that helps to explain this experience:

I wanted to be a circus clown. In order be one I needed to get a job at the circus.

- Someone would want to become a circus clown because they like laughter
- a circus clown sometimes use floppy shoes and a tiny car
- a circus clown has ability to juggle

Teach Open Mind!
Two K-Cap ’03 papers: #2

- Timothy Chklovski: *LEARNER: A System for Acquiring Commonsense Knowledge by Analogy*
Seeded with ~47,000 initial statements from Open Mind common sense

Learner:
- Collected knowledge from > 3400 distinct IP addresses over a 11 month period
- Each such “contributor” averaged ~50 statements
- Total: over 170,000 pieces of knowledge added
Learning new questions

- Given some topic, generate questions that can be used to capture more knowledge

Learning about **NEWSPAPER**

Teach about: newspaper

Examples: beach, chocolate, computer

Similar topics: book, map, magazine, bag

- newspapers contain information?
- all newspapers have pages?
- newspapers are for reading?
Method: Cumulative analogy

- Learner builds a large knowledge base:
  - Objects
  - The properties of an object
  - Truth values \{ T, F \} about statements about objects and their properties

- To suggest new questions about object O:
  - Find similar (analogous) objects
  - Find properties of those objects that are missing in O
Goal is to ask questions about properties that are expected to be hold true

<table>
<thead>
<tr>
<th>Objects</th>
<th>Properties</th>
<th>Has pages</th>
<th>Is for reading</th>
<th>Is cold</th>
<th>Contains knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Ice</td>
<td>?</td>
<td>F</td>
<td>T</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Newspaper</td>
<td>?</td>
<td>T</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Magazine</td>
<td>T</td>
<td>T</td>
<td>?</td>
<td>T</td>
<td></td>
</tr>
</tbody>
</table>
Goal is to ask questions about properties that are expected to be hold true

<table>
<thead>
<tr>
<th>Objects</th>
<th>Properties</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Has pages</td>
<td>Is for reading</td>
<td>Is cold</td>
<td>Contains knowledge</td>
</tr>
<tr>
<td>Book</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>Ice</td>
<td>?</td>
<td>F</td>
<td>T</td>
<td>?</td>
</tr>
<tr>
<td>Newspaper</td>
<td>?</td>
<td>T</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Magazine</td>
<td>T</td>
<td>T</td>
<td>?</td>
<td>T</td>
</tr>
</tbody>
</table>
Results

- How good (likely) are Learner’s questions?
  - From a sample of 1000 questions, 45% were affirmed as true by users
  - If the cumulative analogy method is replaced by a stub – only 8% are affirmed as true

- Cumulative analogy is noise tolerant
Take-home messages

- Actual applications are still toy, but well suited for “fail soft” settings
- Capturing common sense knowledge from web users:
  - Large KBs readily acquired
  - Quality / noise issues can be overcome
Progress since ’03?

- AAAI ’05 papers from earlier today:
  - Searching for common sense: Populating CYC from the web [Matuszek et al]
  - An analysis of knowledge collected from volunteer contributors [Chklovski & Gil]

- Other publications from Singh & Chklovski:
  - AI Mag ’04
  - IUI ’05
  - K-Cap ‘05
Progress since ’03?

- AAAI ’05 papers from earlier today:
  - Searching for Common Sense: Populating CYC from the Web
  - An analysis of knowledge collected from volunteer contributors

- Other publications from Singh & Chklovski:
  - AI Mag ’04
  - IUI ’05
  - K-Cap ‘05

Call for participation! Oct 2-5 in Banff, Canada
See www.kcap05.org
Thank you!