Protégé: Past, Present, and Future

Ray Fergerson
Stanford
Past

• Ancient History (1985-1997)
  – Mark Musen’s Thesis
  – Protégé-II, Protégé/Win
  – “Workshops” 1-2

• Modern Era (1997-2003)
  – Protégé in Java
  – Workshops 3-6
Mark’s Thesis

- PROTÉGÉ system
- Developed as a tool for building knowledge-acquisition tools (a meta-tool) for medical planning applications
- Hardcoded Ontology (or meta-ontology) geared for planning applications
- Written in Lisp, Ran on Lisp machines
- User community: ~1
Protégé-II

- Generic meta-tool
  - No hardcoded ontology
  - No assumptions about “planning”
- Constructed as three distinct applications:
  - ontology editing
  - form customization
  - instance creation
- Written in Objective-C, Lisp…, Ran in NextStep
- User community ~10
• Four guys from four different countries meeting in a pizza restaurant

• Conclusions:
  – The pizza in Italy is really good
  – Too much Chianti makes discussing ontologies difficult
Protégé/Win

• DARPA funding
  – required MS Windows reimplementation

• Protégé/Win
  – Same three applications as Protégé-II
  – Written in C++, ran on Windows 95+

• User community ~100
Decided to re-implement Protégé/Win as a framework into which user-created plug-ins could be added.
- domain-specific slot widgets
- read/write data to other file formats

Decided to adopt Java programming language
• Prototype of new system: Protégé/Java
• Described ability to add tabs
• Subsequently decided…
  – We haven’t gone “far enough” to integrate with other knowledge-base systems such as Ontolingua, CYC, and Loom
  – To adopt Generic Frame Protocol (OKBC) as our knowledge model
  – Throw away previous work and start over
• Protégé-2000 Release 1.0
• Users demonstrated:
  – new domain independent slot widgets
  – tab widgets
• Protégé-2000 1.6 (almost) released
• Free
• Open source
• Growing and diverse user community
  – 1000 registered users
• Protégé-2000 1.9 (almost) released
  – Protégé 2.0 promised for “fall 2003”
    • delivered in February, 2004
• Multiuser client & server
• Reimplemented Diagram/GraphWidget
• OWL Plugin Beta
• Almost 10000 registered users
Present

- Staff
- Mailing Lists
- User community
- Funding
- Release schedule
  - Since last workshop
  - Going forward
Protégé Team

• Research
  – Monica Crubézy
  – Olivier Dameron *new* semantic web tech, modeling principles
  – Holger Knublauch
  – Natasha Noy
  – Daniel Rubin *new* biomedicine, imaging, NLP
  – Samson Tu

• Administration
  – Ted Hopper *new* making sure things run smoothly

• Development
  – Ray Fergerson
  – Jennifer Vendetti
Mailing Lists

- protege-discussion
  - 2300 members
  - 9700 messages
- Created protege-owl mailing list
- Messages viewable as newsgroups, html
  http://gmane.org
Protege Weekly Registration Totals

Total Registered Users: ~20000
## 2004 Conference Attendees

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>75</td>
</tr>
<tr>
<td>Commercial</td>
<td>72</td>
</tr>
<tr>
<td>Government</td>
<td>45</td>
</tr>
<tr>
<td>Non-Profit</td>
<td>17</td>
</tr>
</tbody>
</table>
Conference Attendance

Year

People


0 50 100 150 200 250
Funding

• National Library of Medicine
  – Resource Grant: “National Resource for Biomedical Ontologies and Knowledge Bases”

• National Cancer Institute Center for Bioinformatics
  – Development Contract

• Protege Affiliates:
  – Daimler-Chrysler
Since Last Workshop

- Release 2.0 & 2.1
- Multiuser client & server
- Support for ~5M frames
- Web Browser Interface
- OWL Support
- Improved support for plugins
  - Bundling
  - Isolation
  - Built-in About Box and Documentation support
Future

- Continuing evolution of standard frame-based and OWL Plugin systems
- User-Interface Improvements
- Infrastructure Improvements
- Schedule “Improvements”
Core and OWL Plugin Systems

• OWL Plugin DL support layered on top of the core OKBC frame support
  – Allows conversion between two systems
  – Poses a number of development challenges
  – Provides advantages of giving users:
    • Choice of simpler frame-based or DL interface for editing OWL
    • Access to a variety of plugins

• Plan continued parallel system development and evolution
Retirement Announcement
User Interface Improvements

• Application Changes
  – Cleaner
  – More Attractive
  – More Professional
  – Easier to Learn
  – Easier to Navigate

• Web Site Changes
  “ditto”
Class Browser:
- For Project: Newspaper
- Display: Class Hierarchy by Subclass
- Classes:
  - Chief Honcho
  - Mr. Science
  - Ms. Gardiner
  - Sports Nut

Instance List:
- For Class: Editor
- Display:
  - name (selected)

Class Editor:
- For Class: Editor
- Name: Editor
- Role: Concrete
- Documentation:
  - Editors are responsible for the content of sections
- Constraints:
  - editor-employees-salary-constraint

Template Slots:
- Name: sections
  - Type: Instance
  - Cardinality: multiple
  - Other Facets: classes=[Section]
- responsible_for
  - Type: Instance
  - Cardinality: multiple
  - Other Facets: classes=[Employee]
- name
  - Type: String
  - Cardinality: single
- salary
  - Type: Float
  - Cardinality: single
- date_hired
  - Type: String
  - Cardinality: single
- current_job_title
  - Type: String
  - Cardinality: single
- other_information
  - Type: String
  - Cardinality: single
- phone_number
  - Type: String
  - Cardinality: single
Infrastructure Improvements

- More Flexible Inclusion
  - Extensions of included frames
  - Database inclusion
- Namespaces
  - already in OWL Plugin
- Internationalization Support
  - Localization
  - Alternate names & values
Schedule “Improvements”

- Now aim for two releases/year
  - Before conference (June)
  - ~6 months later
- Betas releases roughly weekly
- Compromise between:
  - “constant updates”
  - “too long before bugs get fixed”
Conference 7 Wish List

• ??? (Come back on Friday at 11:00 am)