NLP at Cambridge

Amad Hussain & Rob LaTour

People

- Prof. Edward J. Briscoe
  Robust parsing technology, constraint-based processing, automated tutoring of second language learners, language acquisition and language evolution.

- Dr. Stephen Clark
  Statistical parsing, compositional and distributional semantics, lexical and world knowledge acquisition, machine learning for NLP.

People Continued

- Prof. Ann Copestake
  Compositional and lexical semantics, formalisms for language representation, multiword expressions, generation, grammar development environments.

- Dr. Simone Teufel
  Text summarisation, text generation and regeneration (sic), information retrieval.

They also have a bunch of Postdocs

Dr. Ronan Cummins
Dr. Mark Granroth-Wilding
Dr. Aurelie Herbelot
Dr. Ekaterina Kochmar
Dr. Tamara Polajnar
Dr. Marek Rei
Dr. Laura Rimell
Dr. Eva Maria Vecchi
Dr. Helen Yannakoudakis

Current/Recent Projects (There are quite a few!)

- The Institute for Automated Language Teaching and Assessment (ALTA)
- Distributional Compositional Semantics for Text Processing (DisCoTex)
- The What-If Machine (WHIM)
- SpaceBook - Spatial & Personal Adaptive Communication Environment
- A Unified Model of Compositional and Distributional Semantics: Theory and Applications
- The Education First-Cambridge Learner Corpus of English - a data driven approach to second language learning
- PANACEA - Platform for Automatic, Normalized Annotation and Cost-Effective Acquisition of Language Resources for Human Language Technologies
- FAUST - Feedback for User Adaptive Statistical Translation
- Computational Natural Language Processing and the Neuro-Cognition of Language
- CRAB: Using Text Mining to Aid Cancer Risk Assessment
- Integrating pragmatic insights with HPSG
- Applying Computational Semantics
- Delph-in interfaces project, funded by Boeing

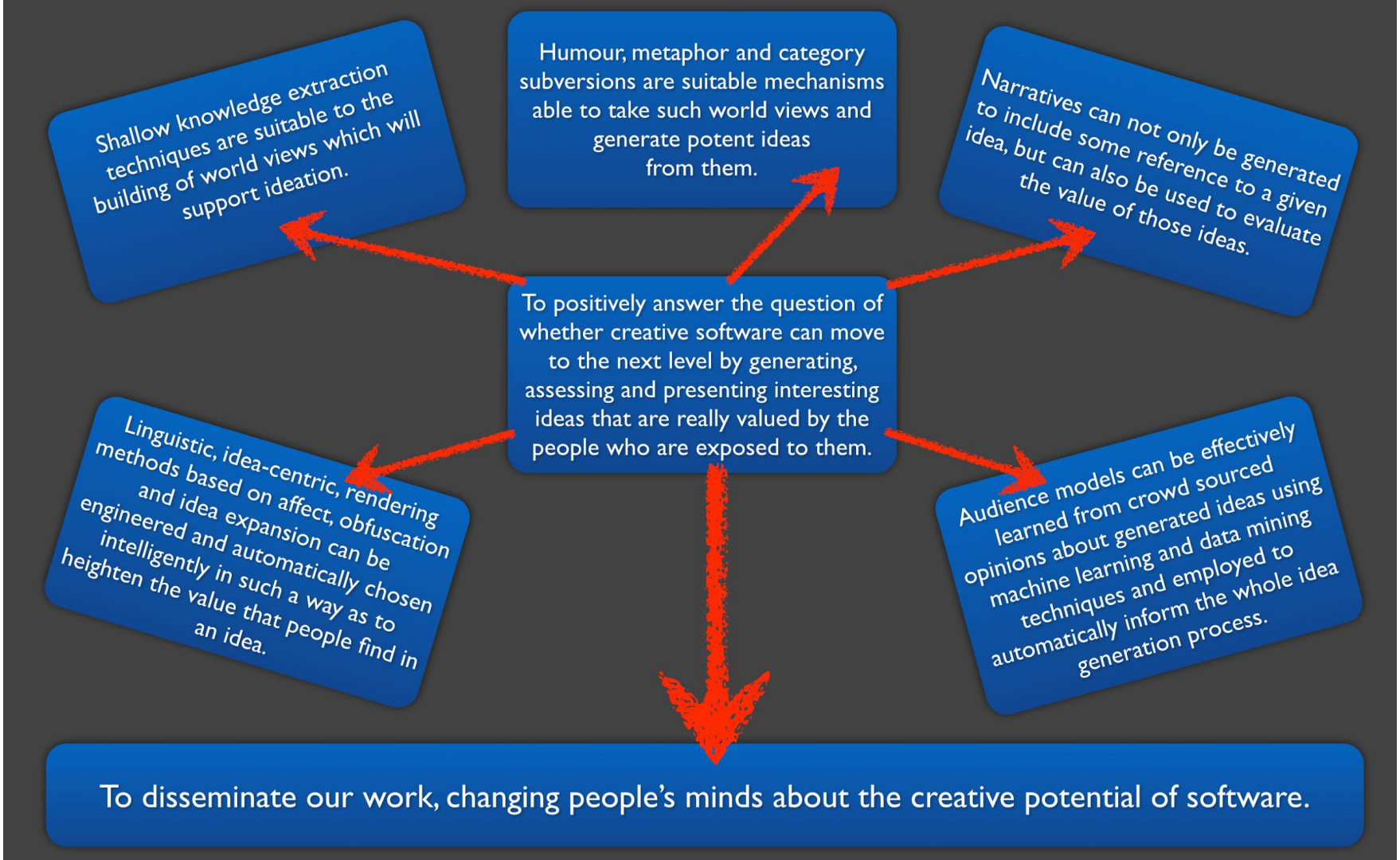
Let's look at a couple: The What-If Machine

Website: http://www.whim-project.eu/whatifmachine/#/welcome

Goal: build a software system able to invent, evaluate and present fictional ideas with real cultural value for artefacts such as stories, jokes, films, paintings and advertisements.

The next slide has a really nice diagram describing the goals of the project.

(We might not be able to play with it, as it didn't seem to be working earlier)


Shallow knowledge extraction techniques are suitable to the building of world views which will support ideation.
Humour, metaphor and category subversions are suitable mechanisms able to take such world views and generate potent ideas from them.
Narratives can not only be generated to include some reference to a given idea, but can also be used to evaluate the value of those ideas.
To positively answer the question of whether creative software can move to the next level by generating, assessing and presenting interesting ideas that are really valued by the people who are exposed to them.
Linguistic, idea-centric, rendering methods based on affect, obfuscation and idea expansion can be engineered and automatically chosen intelligently in such a way as to heighten the value that people find in an idea.
Audience models can be effectively learned from crowd sourced opinions about generated ideas using machine learning and data mining techniques and employed to automatically inform the whole idea generation process.
To disseminate our work, changing people's minds about the creative potential of software.

Let's look at a couple: SpaceBook

Website: http://www.spacebook-project.eu/

Goal: Demonstrate that a personalized speech-only navigational tool is both feasible and practical.

Instead of having to look at a map or ask for directions, you have a verbal interface that gives descriptive and detailed navigational instructions to pedestrians.

Questions?