



WS'03: Preworkshop Activities

Sanjeev Khudanpur

July 14, 2003.

WS'03 Opening Day Presentations

The Annual Workshop Cycle:

It Really Is a Year-Round Activity for CLSP!

September	October	November	December
Call for Proposals	Proposal Deadline Preliminary Screening Invitations to Meeting	Peer Review Meeting Select Team Leaders Form Skeletal Teams	\$\$ from NSF to JHU Advertise for UGs (Deadline Feb 15)
January	February	March	April
Finalize Sr. Members Data Requirements Computing Req's	Select Grad Students Data Contracts Given Computers Ordered	Select Undergrads First Team Meetings Prelim. Data Arrives	Computers Installed Refine Data Specs Perform Pilot Expts
May	June	July	August
2 nd Team Meetings Most Data Arrives Start Baseline Expts	Computing Fine-tuned All the Data Arrives Finish Baseline Expts	Summer School Workshop Begins Team Updates	Workshop Ends Write Final Reports



The Road to the WS'03 Final Four

- Open call for proposals issued Aug 26, 2002.
 - Electronically advertised to 100's of researchers
 - Posted at various conference, e.g. ICSLP in Denver
 - Recruited program managers to help in advertising
- 18 proposals received; some after Oct 15 deadline.
 - Initial screening reduced it to 11 viable proposals
- 2-day peer-review panel met Nov 8-10, 2002.
 - Authors, peers, additional wise-persons, government program managers and JHU faculty were invited
 - Single transferable voting scheme was used
 - 4 highest-rated proposals were selected



The November Meeting Attendees

(and Last-Minute Drop-Outs)

James Baker	Diana Maynard	Tom Crystal	Bill Byrne
Roberto Basili	Franz Och	George Doddington	Jason Eisner
Ciprian Chelba	Misha Pavel	C Van Ess Dykema	Frederick Jelinek
Jordan Cohen	Joseph Picone	Jack Godfrey	Sanjeev Khudanpur
Francesco Danza	Drago Radev	Mary Harper	David Yarowsky
Simona Gandrabur	Brian Roark	Charles Wayne	
Yuqing Gao	John Shore		
Daniel Gildea	Thomas Zheng		
Louise Guthrie	Richard Sproat		

JHU Faculty: 20%

Other Researchers: 60%

Government folks: 20%



Authors (+ Co-Authors) of Proposals Debated on Nov 9-10

- Basili+ : Lexical Boot-strapping: co-training from sparse data
- Chelba : Information Extraction from Natural Language on Relational Domains
- Foster+ : Confidence Estimation for Natural Language Applications
- Gildea : Loosely Tree-based Statistical Machine Translation
- Guthrie+ : Semantic Annotation
 - Och : Syntax for Statistical Machine Translation
- Pavel : Fusion-based Robust Signal Processing for ASR
- Picone : Improving Recognition Performance in Noisy Environment
- Radev : Using the Web for Question Answering
- Sproat+ : Speech Data Mining for National Defense Applications
- Zheng : Dialectal Chinese Speech Recognition Framework



Team Leaders and Projects Selected on November 10, 2002

- Basili+ : Lexical Boot-strapping: co-training from parsing data
- Chelba : Information Extraction from Natural Language on Relational Domains
- Foster+ : Confidence Estimation for Natural Language Applications
- Gildea : Loosely Tree-based Statistical Machine Translation
- Guthrie+ : Semantic Annotation
- Och : Syntax for Statistical Machine Translation
- Pavel : Fusion-based Robust Signal Processing for ASR
- Picone : Improving Recognition Performance in Noisy Environment
- Radev : Using the Web for Question Answering
- Sproat+ : Speech Data Mining for National Defense Applications
- Zheng : Dialectal Chinese Speech Recognition Framework



Team Formation and Spring Planning Meetings

- Team formation process
 - Most senior members of the teams selected over Christmas
 - Graduate students added in Jan-Feb, 2003
 - Undergraduates added in March, 2003
- Fred Jelinek has already covered WS03 demographics
- Each team has had two face-to-face meetings.
 - Syntax for MT: March 15-16 and May 26-27
 - Confidence Annotation: Feb 18 and May 26-27
 - Semantic Annotation: April ?? and June ??
 - Dialectal Mandarin ASR: Jan 18-19 and **SARS!**

That Brings Us to June 30, 2003

Students Arrive in Baltimore!

September	October	November	December
Call for Proposals	Proposal Deadline Preliminary Screening Invitations to Meeting	Peer Review Meeting Select Team Leaders Form Skeletal Teams	Advertise for UGs (Deadline Feb 15)
January	February	March	April
Finalize Sr. Members Data Spec's Computing Req's	Select Grad Students 1 Data Contract Given	Select Undergrads First Team Meetings	Xeon Computers Add'l Data Contract
May	June	July	August
2 nd Team Meetings Start Baseline Expts \$\$ from NSF to JHU	Finish Baseline Expts?	Summer School Workshop Begins Team Updates	Workshop Ends Write Final Reports



The JHU Summer School on Human Language Technology

- A two week tutorial on all aspects of human language technology by expert academics and practitioners.
 - ASR, TTS, IR, IE, summarization, generation, MT.
- Started in 1998 to bring workshop undergraduate students up to speed on human language technology.
 - Attended by DoD personnel looking for a short comprehensive overview of language technology
 - Additional participants starting 2002 from NAACL
- The attendance keeps growing!
 - 10 undergraduates (8 WS02, 2 NAACL)
 - 22 graduate students (10 WS02, 8 NAACL, 4 Others)
 - 5-6 researchers (NSA, ARL, other government)

Summer School: 2003 Edition

<http://www.clsp.jhu.edu/ws03/calendar/preliminary.shtml>

Mon June 30 NLP	Tue July 1 Speech Proc	Wed July 2 IR and CLIR	Thu July 3 IR and CLIR	Fri July 4 Independence Day Holiday
C Manning	Bill Byrne	Jim Mayfield	Regina Barzilay	
Jason Eisner	Dan Ellis	David Yarowsky	Drago Radev	
Jason Eisner Noah Smith	Dan Ellis V Doumptotis & Y Deng	Philip Resnik Elliott Drabek & Y Deng	Drago Radev Elliott Drabek	G Ambwani & Yi Su A Emami & S Kumar

Mon July 7 Dialog Sys	Tue July 8 Machine Learn	Wed July 9 Syntax for MT	Thu July 10 Confidence Est	Fri July 11 Semantic Annot
S Seneff	Ray Mooney	Kenji Yamada	Yoshua Bengio	Ralph Grishman
Grace Chung		Salim Roukos	George Foster	Doug Appelt
Grace Chung Seneff & J Polifroni	Jeff Bilmes Ahmad Emami	Anoop Sarkar S Kumar & Alex Fraser	S. Gandrabur Erin Fitzgerald	Louise Guthrie D Guthrie & J Cui

The Workshop Calendar

<http://www.clsp.jhu.edu/ws03/calendar>

Mon Jul 14	Tue Jul 15	Wed Jul 16	Thu Jul 17	Fri Jul 18
Opening Presentation Steering Comm Mtg		Jim West Seminar	UG Lunch Group Dinner	
Mon Jul 21	Tue Jul 22	Wed Jul 23	Thu Jul 24	Fri Jul 25
Steering Comm Mtg		Herman Ney Seminar Team Updates	UG Lunch Group Dinner	Townhall Meeting
Mon Jul 28	Tue Jul 29	Wed Jul 30	Thu Jul 31	Fri Aug 1
Steering Comm Mtg		Seminar TBA	UG Lunch Group Dinner	
Mon Aug 4	Tue Aug 4	Wed Aug 6	Thu Aug 7	Fri Aug 8
Steering Comm Mtg		Seminar TBA Team Updates	UG Lunch Group Dinner	Townhall Meeting
Mon Aug 11	Tue Aug 12	Wed Aug 13	Thu Aug 14	Fri Aug 15
Steering Comm Mtg		Seminar TBA	UG Lunch Group Dinner	
Mon Aug 18	Tue Aug 19	Wed Aug 20	Thu Aug 21	Fri Aug 22
Steering Comm Mtg		Final Presentation	Final Presentation Dinner Party	Lab Clean-Up Day Farewell Tea