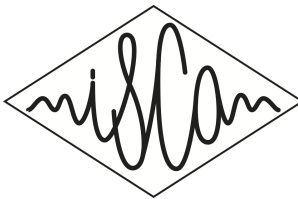SIGDIAL 2012

13th Annual Meeting of the Special Interest Group on Discourse and Dialogue

Conference Handbook

5-6 July 2012
Seoul National University
Seoul, South Korea

Contents

Introduction

It is our great pleasure to welcome you to the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL). The conference is held in Seoul, South Korea, July 5-6, 2012, in affiliation with the 50th Annual Meeting of the ACL.

SIGDIAL continues to serve as a publication venue for research that spans many aspects of discourse and dialogue. The special theme for SIGDIAL 2012 is "characterizing dialog coherence". The program includes two keynotes, eighteen oral presentations, twenty posters and three demos. We are deeply honored to have the participation of our two keynote speakers: Professor Tatsuya Kawahara (Kyoto University) and Professor Diane Litman (University of Pittsburgh). We thank the members of the Program Committee, who provided excellent reviews. Special thanks go to Giuseppe Di Fabbrizio and Christophe Cerisara for providing last minute reviews.

We thank Kallirroi Georgila, Mentoring Chair, for coordinating the paper mentoring process in the limited time between notification and final submission of papers. This year, nine of the accepted papers were mentored. We thank the mentors: Ron Artstein, Srinivas Bangalore, Michael Johnston, Kristiina Jokinen, Vincent Ng, Andrei Popescu-Belis, David Schlangen, Amanda Stent, and Michael Strube.

We thank Minhwa Chung, Local Arrangements Chair, and Hyung Soon Kim, Jungyun Seo and Sunhee Kim, the members of the Local Arrangements Committee, for taking care of the many details of the local arrangements. We also thank the student volunteers for on-site assistance: Minsoo Na, HyukSu Ryu, Hyejin Hong, Yumi Hwang, Hyosung Hwang, Jieun Song, Ji Yeon Lee, Won-Seok Choi, Byul Park, Xiang-Lan Xu, Ethan Selfridge and William Wang.

We thank Jason Williams, Sponsorship Chair, for recruiting and liaising with our conference sponsors. Sponsorship makes possible valuable aspects of the SIGDIAL program, such as the invited speakers, conference dinner, and best paper awards. We gratefully acknowledge the support of our sponsors, including AT&T, AVIOS, Honda Research Institute, IBM Research, KT Corporation, Microsoft Research, NHN Corporation, and Seoul National University.

We thank last year's Program Co-Chairs, Joyce Chai and Rebecca Passonneau, last year's General Co-Chairs, Johanna Moore and David Traum, and last year's Local Chair, Peter Heeman, for their helpful advice and answers to questions. We also thank the SIGDIAL board, in particular Tim Paek, Amanda Stent, and Kristiina Jokinen, for their advice and support.

We thank Priscilla Rasmussen at the ACL for handling the financial transactions for SIGDIAL 2012, including advance registration.

We thank SoftConf and the ACL for use of the START conference management system.

Finally, we thank all the authors of the papers in this volume, and all the conference participants for making this event such a great opportunity for new research in discourse and discourse.

Gary Geunbae Lee and Jonathan Ginzburg
General Co-Chairs

Claire Gardent and Amanda Stent
Technical Program Co-Chairs

Conference Organization

General Co-Chairs:

Gary Geunbae Lee, POSTECH
Jonathan Ginzburg, Université Paris-Diderot

Program Co-Chairs:

Claire Gardent, CNRS/LORIA Nancy
Amanda Stent, AT&T Labs - Research

Local Chair:

Minhwa Chung, Seoul National University

Local Arrangements Committee:

Hyung Soon Kim, Pusan National University (KSSS liaison)
Jungyun Seo, Sogang University (KIISE liaison)
Sunhee Kim, Seoul National University (SNU liaison)

Mentoring Chair:

Kallirroi Georgila, University of Southern California Institute for Creative Technologies (ICT)

Sponsorships Chair:

Jason D. Williams, Microsoft Research

SIGDIAL Officers:

President: Tim Paek, Microsoft Research
Vice President: Amanda Stent, AT&T Labs - Research
Secretary/Treasurer: Kristiina Jokinen, University of Helsinki

Program Committee:

Hua Ai, Georgia Institute of Technology
Jan Alexandersson, German Research Institute for Artificial Intelligence (DFKI GmbH)
Ron Artstein, University of Southern California Institute for Creative Technologies (ICT)
Srinivas Bangalore, AT&T Labs Research
Luciana Benotti, Universidad Nacional de Cordoba
Nate Blaylock, Florida Institute for Human Machine Cognition (IHMC)
Dan Bohus, Microsoft Research
Johan Boye, KTH Royal Institute of Technology
Kristy Boyer, North Carolina State University
Harry Bunt, Tilburg University
Sandra Carberry, University of Delaware

Jennifer Chu-Carroll, IBM T. J. Watson Research Center
Mark Core, University of Southern California Institute for Creative Technologies (ICT)
Heriberto Cuayáhuitl, German Research Institute for Artificial Intelligence (DFKI GmbH)
David DeVault, University of Southern California Institute for Creative Technologies (ICT)
Barbara Di Eugenio, University of Illinois at Chicago
Myroslava Dzikovksa, University of Edinburgh
Maxine Eskenazi, Carnegie Mellon University
Raquel Fernández Rovira, University of Amsterdam
Katherine Forbes-Riley, University of Pittsburgh
Matthew Frampton, Institute of Cancer Research
Kallirroi Georgila, University of Southern California Institute for Creative Technologies (ICT)
Agustin Gravano, University of Buenos Aires
Nancy Green, University of North Carolina at Greensboro
Alexander Gruenstein, Google
Curry Guinn, University of North Carolina at Wilmington
Joakim Gustafson, KTH Royal Institute of Technology
Peter Heeman, Oregon Health & Science University
Julia Hirschberg, Columbia University
Michael Johnston, AT&T Labs - Research
Kristiina Jokinen, University of Helsinki
Arne Jönsson, Linköping University
Pamela Jordan, University of Pittsburgh
Simon Keizer, Heriot-Watt University
Alistair Knott, University of Otago
Kazunori Komatani, Nagoya University
Stefan Kopp, Bielefeld University
Ian Lane, Carnegie Mellon University
Sungjin Lee, POSTECH
Oliver Lemon, Heriot-Watt University
James Lester, North Carolina State University
Anton Leuski, University of Southern California Institute for Creative Technologies (ICT)
Diane Litman, University of Pittsburgh
Ramón López-Cózar, University of Granada
Wolfgang Minker, University of Ulm
Sebastian Möller, Telekom Innovation Laboratories, TU Berlin
Mikio Nakano, Honda Research Institute
Ani Nenkova, University of Pennsylvania
Vincent Ng, University of Texas at Dallas
Tim Paek, Microsoft Research
Olivier Pietquin, Supélec - Metz Campus, UMI 2958 (GeorgiaTech - CNRS)
Paul Piwek, Open University
Andrei Popescu-Belis, Idiap Research Institute
Matthew Purver, Queen Mary University of London
Antoine Raux, Honda Research Institute
Norbert Reithinger, German Research Institute for Artificial Intelligence (DFKI GmbH)
Hannes Rieser, Bielefeld University
Carolyn Penstein Rosé, Carnegie Mellon University
Alex Rudnicky, Carnegie Mellon University

David Schlangen, Bielefeld University
Elizabeth Shriberg, Microsoft Research
Candace Sidner, Worcester Polytechnic Institute
Gabriel Skantze, KTH Royal Institute of Technology
Manfred Stede, University of Potsdam
Svetlana Stoyanchev, Columbia University
Michael Strube, Heidelberg Institute for Theoretical Studies
Marc Swerts, Tilburg University
Takenobu Tokunaga, Tokyo Institute of Technology
Thora Tenbrink, University of Bremen
Joel Tetreault, Educational Testing Service
Gokhan Tur, Microsoft Research
Marilyn Walker, University of California at Santa Cruz
Nigel Ward, University of Texas El Paso
Jason D. Williams, Microsoft Research
Ingrid Zukerman, Monash University

Additional Reviewers:

Timo Baumann, University of Hamburg
Christoph Broschinski, Bielefeld University
Hendrik Buschmeier, Bielefeld University
Christophe Cerisara, CNRS/LORIA Nancy
Nina Dethlefs, Heriot-Watt University
Giuseppe Di Fabbrizio, AT&T Labs - Research
Ulle Endriss, University of Amsterdam
Rivka Levitan, Columbia University
Alejandra Lorenzo, CNRS/LORIA Nancy

Student Volunteers:

Minsoo Na, Seoul National University
HyukSu Ryu, Seoul National University
Hyejin Hong, Seoul National University
Yumi Hwang, Seoul National University
Hyosung Hwang, Seoul National University
Jieun Song, Seoul National University
Ji Yeon Lee, Seoul National University
Won-Seok Choi, Sogang University
Byul Park, Sogang University
Xiang-Lan Xu, Sogang University
Ethan Selfridge, Oregon Health & Science University
William Wang, Carnegie Mellon University

Invited Speakers:

Tatsuya Kawahara, Kyoto University
Diane Litman, University of Pittsburgh

Video Recording

This year, for the first time, we will be video recording oral presentations at SIGDIAL and making them freely available over the web. The aims of offering video recording are to increase the overall impact and reach of the SIGDIAL Conference, and to provide additional benefit to speakers by creating a new channel for dissemination of their work. After recording and editing, the talks will be hosted by `http://www.superlectures.com`.

We thank Jason Williams for organizing this new initiative, and superlectures.com for hosting the presentations.

Video recording of oral presentations is opt-in. If you are giving an oral presentation at SIGDIAL we hope you will participate in video recording. You will need to provide a signed consent form, and a copy of your slides (PDF or powerpoint). You can provide both of these to Jason Williams.
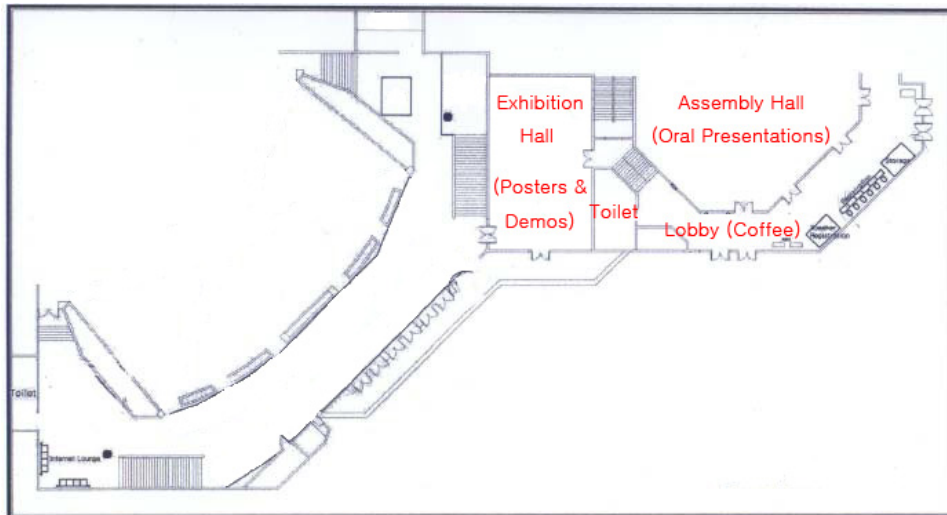
Conference Venue

Figure 1: Floor plan of the Culture Center

SIGDIAL 2012 is held in the Culture Center (Building 73) on the campus of Seoul National University. The oral presentations will be held in the Assembly Hall, and the poster and demo sessions and coffee breaks in the Lobby (see Figure 1).
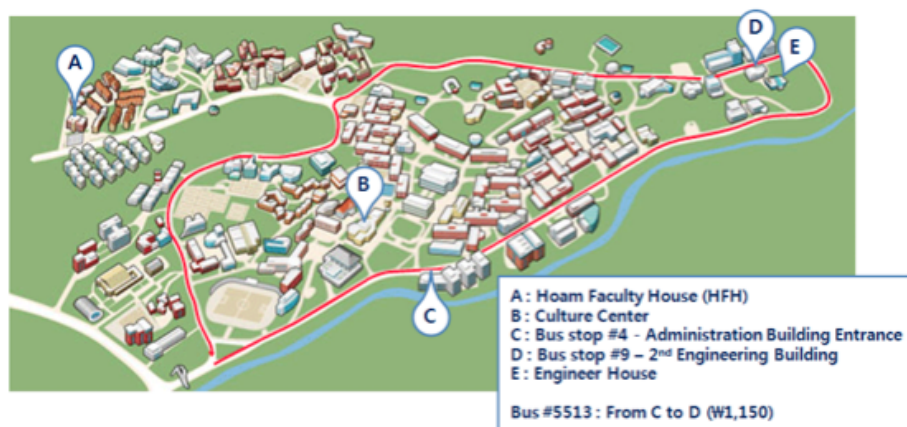

Figure 2: Seoul National University

The conference dinner will be held in the Auditorium (Room 101) of Engineer House (Building 310) on the campus of Seoul National University (see Figure 2). To get there, take bus no. 5513 at bus stop no. 4 (Administration Building Entrance) and get off at bus stop no. 9 (Engineering Building).

Figure 3: Area map

For a detailed map of the university, see `http://www.sigdial.org/workshops/conference13/venue/guidemap_eng_cultural_center.pdf`.

The map in Figure 3 shows the relative locations of the conference venue, hotels and public transit.

How to Get to the Conference Venue

To get to the conference venue, from Subway Line No.2 (Green Line), Seoul National University Entrance Station (Exit No.3):

1. After leaving from exit no.3, walk towards the Gwanak-gu Office.

2. Take the no. 5513 bus or a campus shuttle bus.

3. Get off at the SNU Administration Building (Building 60) stop (6th stop from the Subway station).

4. Walk to Assembly Hall (JungGangDang in Korean) in the Culture Center (Building 73), which is located 200 m from the bus stop.

Caution 1: The SNU Culture Center is 3km from the SNU Entrance station (about 50m on foot). It is recommended to take a bus or a taxi to get the university.
Caution 2: Be sure not to confuse "Seoul National University Entrance Station" with "Seoul National University of Education Station". See **Getting Around**, below.

Getting Around

Seoul has a well-organized public transportation system; you can get to Seoul National University (SNU) easily by using the bus and subway.

Subway

When using the subway, take Subway Line No. 2 (Green Line) and stop at the Seoul National University Entrance Station, the Naksungdae Station, or the Shillim Station.

Subway Line No. 2 (Green Line), SNU Entrance Station (Exit No. 3)

Even though SNU is near the Seoul National University Entrance Station, **it is not within walking distance**. You will need to take a bus or a taxi to get to the university. Be sure not to confuse the SNU Entrance Station with the Seoul National University of Education/Court or Prosecutor's Office Station.

After leaving the SNU Entrance Station (exit no. 3), walk towards the Gwanak-gu Office. There, you will find SNUl shuttle buses and also regular buses that come to SNU. The campus shuttles are free of charge. The operation of the buses is usually in the range of every 5 to 20 minutes. Buses number 5511 and 5513 circulate the campus counter-clockwise, from the university's main entrance to the Administration Building.

Figure 4: Subway Line No. 2 (Green Line), SNU Entrance Station (Exit No. 3)

Subway Line No. 2 (Green Line), Naksungdae Station (Exit No. 4)

You will need to take additional public transportation from the Naksungdae Station. After leaving the station (exit no. 4), walk straight and make a left turn at the gas station. There will be a bus stop in front of the bakery. Take the small green town bus no. 2.
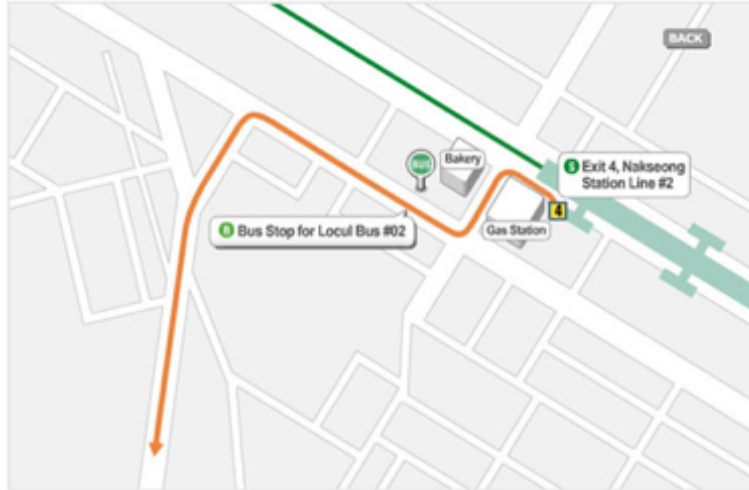

Figure 5: Subway Line No. 2 (Green Line), Naksungdae Station (Exit No. 4)

Subway Line No. 2 (Green Line), Shillim Station (Exit No. 3)

You will need to take additional public transportation from the Shillim Station. After leaving the station (exit no. 3), take bus no. 5516, which circulates counter-clockwise from the main entrance towards the Administration Building.

Figure 6: Subway Line No. 2 (Green Line), Shillim Station (Exit No. 3)

Number	Final Stop	Major Stops
5411(GREEN)	SNU ↔ Yeomgok-dong	SNU Subway Station, Yangjae, Yeomgok
5412(GREEN)	SNU ↔ Kangnam Subway Station	Sadang Station, Express Bus Terminal
5517(GREEN)	SNU ↔ Joong-Ang University	Shillim Station, Noryangjin Station
5522(GREEN)	SNU ↔ Nangok	Shillim, Nangok
5528(GREEN)	SNU ↔ Guro Digital Complex	Shillim, Siheung, Gasan
5613(GREEN)	SNU ↔ Yeouido	Shillim, Daebang
5614(GREEN)	SNU ↔ Kemok Girl's Highschool	Shillim, Daebang
6511(GREEN)	SNU ↔ Guro-dong	Shindaebang Station, Shindorim Station
6512(GREEN)	SNU ↔ Guro-dong	SNU Subway Station, Shillim Station, Youngdeungpo
6513(GREEN)	SNU ↔ Chulsan-dong	Shillim Station, Daebang Station, Youngdeungpo
6514(GREEN)	SNU ↔ Shinjung-dong	Shillim Station, Daebang Station, Youngdeungpo Station, Dangsan Station
501 (BLUE)	SNU ↔ Jongro 2nd district	Sangdo-dong, Yongsan, Seoul Station
502 (BLUE)	SNU ↔ Dongdaemoon Stadium	Shillim, Noryangjin, Hanganjin
750 (BLUE)	SNU ↔ Susaek	Gwanak-gu Office, Sangdo Tunnel, Seoul Station, Shinchon

Table 1: Buses that travel past SNU

Bus

Table 1 lists buses that travel past SNU[1].

Campus Shuttle

Seoul National University provides free shuttle buses for the convenience of its students at Gwanak Campus. There are four types: Off-Campus Shuttle Bus, On-Campus Circulation Shuttle Bus, Library Shuttle Bus, and shuttle buses for students with disabilities.

Off-Campus Shuttle Bus

Off-campus shuttle buses connect Gwanak Campus with SNU subway station line no. 2 and Shillim-dong, where a number of SNU students reside.

Off-campus shuttle buses run from 07:00 to 18:30 every Monday through Friday. This time schedule only applies during the academic year. New time schedules for Summer/Winter Sessions are announced at those times. If you are seeking further detailed information on the above-mentioned locations and time schedules, please refer to the file schedules, please refer to `http://en.snu.ac.kr/upload/about_speech/offcampus_shuttle.pdf`.

SNU subway station ↔ Gwanak Campus

A shuttle bus stop at SNU subway station line no. 2 is located about 100m from exit no. 3. The bus from this stop will take you to the Administration Building (Building 60) at Gwanak Campus. In order to return to the subway station exit no. 2, board the shuttle bus at the same spot you got off (bus stop 4-1, on the right side of the two in Figure 7).

Another shuttle bus at SNU subway station line no. 2 goes to the Engineering Education Building (Buildings 301 and 302). The bus stop is located about 50m from exit no. 3. If you

---

[1] A detailed service route of each bus is provided on the Seoul bus network online service website: `http://bus.seoul.go.kr`

want to go to the subway station from Buildings 301 and 302, you have to take a local city bus (green bus) or take the shuttle bus from the Administration Building (Building 60).
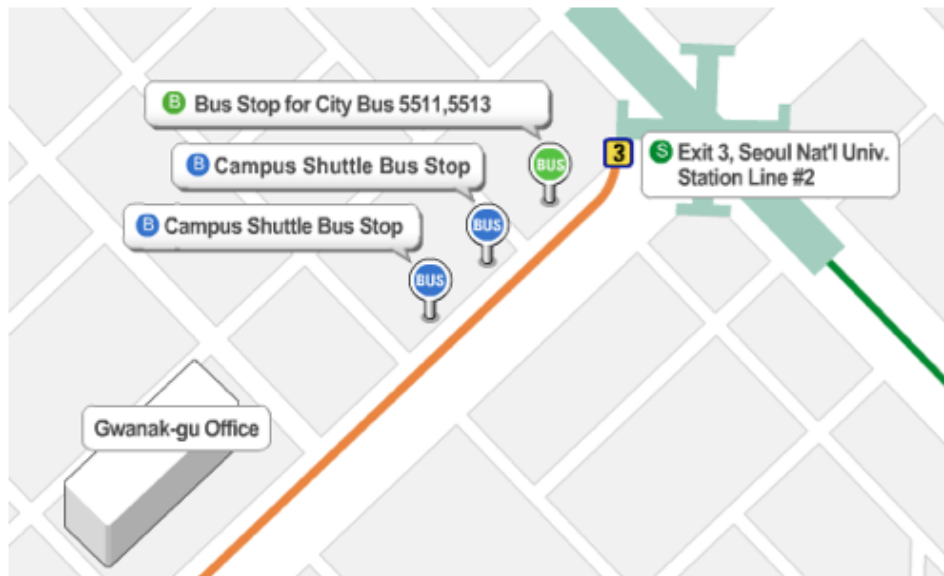

Figure 7: SNU subway station ↔ Gwanak Campus

Shillim-dong ↔ Gwanak Campus

You can take a shuttle bus at Shillim-dong in front of the convenience store "Buy the way" (about 200m from Woori Bank). This bus goes to the Administration Building (Building 60). Boarding the shuttle bus at bus stop 4-1 (on the left side of the two in Figure 7) will take you back to the "Buy the way" store.

On-Campus Circulation Shuttle Bus

On campus, you may choose to use on-campus circulation shuttle buses. Many students take this shuttle bus after arriving at the main gate of SNU by city bus or airplane limousine bus.

There are two on-campus routes. One circulates the whole campus counterclockwise. This route includes 18 bus stops starting from the main gate of SNU and ending at the College of Business Administration (Building 58). During the academic year, this shuttle bus runs from 08:00 to 19:00; it ends one hour earlier during Summer/Winter Sessions. The second route is shorter. It starts from the Administration Building (Building 60) and mainly visits the Engineering Building (Building 301 and 302). The time schedule for this route is from 10:00 to 19:00 during the academic year only. For further detailed information, please refer to the On-Campus Circulation Shuttle Bus Routes:

- http://en.snu.ac.kr/upload/about_speech/shuttlebus_route1.pdf

- http://en.snu.ac.kr/upload/about_speech/shuttlebus_route2.pdf

Local Information

Time Zone

Standard Time in Korea is 9 hours ahead of Greenwich Mean Time (GMT+9).

Banking & ATM

There are plenty of places in Korea to change currencies, including banks, authorized currency exchanges and major hotels. ATM machines are ubiquitous, although not all offer services for foreign cards. Cards with the Plus and Cirrus logos are the easiest to use and most widely accepted in Korea. Banks are open from 9:00 a.m. to 4:00 p.m., Mondays to Fridays.

Currency & Exchange

The unit of currency in Korea is the Korean Won (KRW). Various notes and coins are used. Notes occur in 1,000, 5,000, 10,000 and 50,000 won denominations, while coins occur in 10, 50, 100, and 500 won denominations. As of April 2010, the exchange rate is approximately KRW 1,100 to USD 1 and KRW 1,500 to EUR 1. Foreign bank notes and traveler's checks can be converted into the Korean won at foreign exchange banks and other authorized money exchange outlets. The exchange rate is subject to market fluctuations.

Business Hours

Government office hours are usually from 9:00 to 18:00 on weekdays and are closed on weekends. Most stores are open every day from 10:30 to 20:00, including Sundays.

Tax & VAT

Foreign tourists can receive nearly 10 percent VAT refunds for purchases at duty free shopping outlets. Goods must be taken out of Korea within three months of purchase to be eligible for a tax refund. A Value-Added Tax (VAT) is levied on most goods and services at a standard rate of 10% and is included in the retail price. At hotels, a 10% VAT is applied to rooms, meals and other services and is included in the bill.

Tipping

Tipping is not a customary practice in Korea.

Electricity

The standard voltage in Korea is 220 volts. The outlet has two round holes and is the same type used in France, Germany, Austria, Greece, Turkey, and many other countries.

BBB (Before Babel Brigade): Volunteer Interpretation Service

Whenever and wherever you encounter communication problems with Koreans, just call 1588-5644 and press a number for your preferred language. Your call will be quickly connected to a BBB volunteer's mobile phone through an automated switchboard, and you will soon receive interpretation service.

- BBB volunteer services are available in 16 languages.

- For more information, please refer to the BBB Korea official webpage (`http://www.bbbkorea.org`). If you don't have a mobile phone in Korea, just show a BBB card to any Korean nearby, so that he/she can Help you with their mobile phone.

Conference Program

Thursday July 5, 2012

> **9:00 AM Welcome and Opening Remarks**

> **9:15 AM Keynote 1 (Session Chair: Alex Rudnicky)**

9:15-10:15   *Multi-Modal Sensing and Analysis of Poster Conversations: Toward Smart Poster-board*
Tatsuya Kawahara

> **10:15 AM Coffee Break**

> **10:45 AM Oral Presentation Session 1: Evaluation (Session Chair: Diane Litman)**

10:45–11:10   *An End-to-End Evaluation of Two Situated Dialog Systems*
Lina M. Rojas Barahona, Alejandra Lorenzo and Claire Gardent

11:10–11:35   *"Love ya, jerkface": Using Sparse Log-Linear Models to Build Positive (and Impolite) Relationships with Teens*
William Yang Wang, Samantha Finkelstein, Amy Ogan, Alan W. Black and Justine Cassell

11:35–12:00   *Enhancing Referential Success by Tracking Hearer Gaze*
Alexander Koller, Konstantina Garoufi, Maria Staudte and Matthew Crocker

> **12:00 PM Lunch**

> **1:00 PM Oral Presentation Session 2 (Theme Session): Coherence (Session Chair: David DeVault)**

1:00–1:25   *Unsupervised Topic Modeling Approaches to Decision Summarization in Spoken Meetings*
Lu Wang and Claire Cardie

1:25–1:50   *An Unsupervised Approach to User Simulation: Toward Self-Improving Dialog Systems*
Sungjin Lee and Maxine Eskenazi

1:50–2:15   *Hierarchical Conversation Structure Prediction in Multi-Party Chat*
Elijah Mayfield, David Adamson and Carolyn Penstein Rosé

Thursday July 5, 2012 (continued)

2:15 PM Poster Session 1 "Madness" (Session Chair: David Schlangen)

2:30 PM Poster Session 1

2:30–4:00   *Rapid Development Process of Spoken Dialogue Systems using Collaboratively Constructed Semantic Resources*
Masahiro Araki

2:30–4:00   *The Effect of Cognitive Load on a Statistical Dialogue System*
Milica Gašić, Pirros Tsiakoulis, Matthew Henderson, Blaise Thomson, Kai Yu, Eli Tzirkel and Steve Young

2:30–4:00   *Predicting Adherence to Treatment for Schizophrenia from Dialogue Transcripts*
Christine Howes, Matthew Purver, Rose McCabe, Patrick G. T. Healey and Mary Lavelle

2:30–4:00   *Reinforcement Learning of Question-Answering Dialogue Policies for Virtual Museum Guides*
Teruhisa Misu, Kallirroi Georgila, Anton Leuski and David Traum

2:30–4:00   *From Strangers to Partners: Examining Convergence within a Longitudinal Study of Task-Oriented Dialogue*
Christopher M. Mitchell, Kristy Elizabeth Boyer and James C. Lester

2:30-4:00   *The Structure and Generality of Spoken Route Instructions*
Aasish Pappu and Alexander Rudnicky

2:30–4:00   *Improving Implicit Discourse Relation Recognition Through Feature Set Optimization*
Joonsuk Park and Claire Cardie

2:30–4:00   *A Temporal Simulator for Developing Turn-Taking Methods for Spoken Dialogue Systems*
Ethan O. Selfridge and Peter A. Heeman

2:30–4:00   *Dialogue Act Recognition using Reweighted Speaker Adaptation*
Congkai Sun and Louis-Philippe Morency

2:30–4:00   *Estimating Adaptation of Dialogue Partners with Different Verbal Intelligence*
Kseniya Zablotskaya, Fernando Fernández-Martínez and Wolfgang Minker

Thursday July 5, 2012 (continued)

2:30 PM Demo Session

2:30–4:00   *A Demonstration of Incremental Speech Understanding and Confidence Estimation in a Virtual Human Dialogue System*
David DeVault and David Traum

2:30–4:00   *Integrating Location, Visibility, and Question-Answering in a Spoken Dialogue System for Pedestrian City Exploration*
Srinivasan Janarthanam, Oliver Lemon, Xingkun Liu, Phil Bartie, William Mackaness, Tiphaine Dalmas and Jana Goetze

2:30–4:00   *A Mixed-Initiative Conversational Dialogue System for Healthcare*
Fabrizio Morbini, Eric Forbell, David DeVault, Kenji Sagae, David Traum and Albert Rizzo

4:00 PM Sponsor Presentations 1 (Jason D. Williams)

4:25 PM Oral Presentation Session 3: Discourse Structure (Session Chair: Amanda Stent)

4:25–4:50   *Towards Mediating Shared Perceptual Basis in Situated Dialogue*
Changsong Liu, Rui Fang and Joyce Chai

4:50–5:15   *Global Features for Shallow Discourse Parsing*
Sucheta Ghosh, Giuseppe Riccardi and Richard Johansson

5:15–5:40   *A Reranking Model for Discourse Segmentation using Subtree Features*
Ngo Xuan Bach, Nguyen Le Minh and Akira Shimazu

Conference Reception and Dinner at Engineer House, SNU

Friday July 6, 2012

9:00 AM Announcements

9:15 AM Oral Presentation Session 4: Statistical Models of Dialog (Session Chair: Gary Geunbae Lee)

9:15–9:40   *Landmark-Based Location Belief Tracking in a Spoken Dialog System*
Yi Ma, Antoine Raux, Deepak Ramachandran and Rakesh Gupta

9:40–10:05   *Probabilistic Dialogue Models with Prior Domain Knowledge*
Pierre Lison

10:05–10:30   *Exploiting Machine-Transcribed Dialog Corpus to Improve Multiple Dialog States Tracking Methods*
Sungjin Lee and Maxine Eskenazi

10:30 AM Coffee Break

11:00 AM Keynote 2 (Session Chair: David Traum)

11:00–12:00   *Cohesion, Entrainment and Task Success in Educational Dialog*
Diane Litman

12:00 PM Lunch and Business Meeting

1:30 PM Oral Presentation Session 5: Paralinguistic Features (Session Chair: Tatsuya Kawahara)

1:30–1:55   *A Bottom-Up Exploration of the Dimensions of Dialog State in Spoken Interaction*
Nigel G. Ward and Alejandro Vega

1:55–2:20   *Using Group History to Identify Character-Directed Utterances in Multi-Child Interactions*
Hannaneh Hajishirzi, Jill F. Lehman and Jessica K. Hodgins

2:20–2:45   *Adapting to Multiple Affective States in Spoken Dialogue*
Kate Forbes-Riley and Diane Litman

Friday July 6, 2012 (continued)

2:45 PM Poster 2 "Madness" (Session Chair: Antoine Raux)

3:00 PM Poster Session 2

3:00–4:00   *Dialog System Using Real-Time Crowdsourcing and Twitter Large-Scale Corpus*
Fumihiro Bessho, Tatsuya Harada and Yasuo Kuniyoshi

3:00–4:00   *Automatically Acquiring Fine-Grained Information Status Distinctions in German*
Aoife Cahill and Arndt Riester

3:00–4:00   *A Unified Probabilistic Approach to Referring Expressions*
Kotaro Funakoshi, Mikio Nakano, Takenobu Tokunaga and Ryu Iida

3:00–4:00   *Combining Verbal and Nonverbal Features to Overcome the "Information Gap" in Task-Oriented Dialogue*
Eun Young Ha, Joseph F. Grafsgaard, Christopher M. Mitchell, Kristy Elizabeth Boyer and James C. Lester

3:00–4:00   *Semantic Specificity in Spoken Dialogue Requests*
Ben Hixon, Rebecca J. Passonneau and Susan L. Epstein

3:00–4:00   *Contingency and Comparison Relation Labeling and Structure Prediction in Chinese Sentences*
Hen-Hsen Huang and Hsin-Hsi Chen

3:00–4:00   *A Study in How NLU Performance Can Affect the Choice of Dialogue System Architecture*
Anton Leuski and David DeVault

3:00–4:00   *Integrating Incremental Speech Recognition and POMDP-Based Dialogue Systems*
Ethan O. Selfridge, Iker Arizmendi, Peter A. Heeman and Jason D. Williams

3:00–4:00   *A Regression-Based Approach to Modeling Addressee Backchannels*
Allison Terrell and Bilge Mutlu

3:00–4:00   *Improving Sentence Completion in Dialogues with Multi-Modal Features*
Anruo Wang, Barbara Di Eugenio and Lin Chen

Friday July 6, 2012 (continued)

4:00 PM Sponsor Presentations 2 (Jason D. Williams)

4:25 PM Oral Presentation Session 6: NLG and NLU (Session Chair: Minhwa Chung)

4:25–4:50 *Combining Incremental Language Generation and Incremental Speech Synthesis for Adaptive Information Presentation*
Hendrik Buschmeier, Timo Baumann, Benjamin Dosch, Stefan Kopp and David Schlangen

4:50–5:15 *Focused Meeting Summarization via Unsupervised Relation Extraction*
Lu Wang and Claire Cardie

5:15–5:40 *Markov Logic Networks for Situated Incremental Natural Language Understanding*
Casey Kennington and David Schlangen

5:40 PM Best Paper Awards and Concluding Remarks

Conference Abstracts

Multi-modal Sensing and Analysis of Poster Conversations: Toward Smart Posterboard
Tatsuya Kawahara
Thursday July 5, 2012 9:15-10:15

Conversations in poster sessions in academic events, referred to as poster conversations, pose interesting and challenging topics on multi-modal analysis of multi-party dialogue. This article gives an overview of our project on multi-modal sensing, analysis and "understanding" of poster conversations. We focus on the audience's feedback behaviors such as non-lexical backchannels (reactive tokens) and noddings as well as joint eye-gaze events by the presenter and the audience. We investigate whether we can predict when and who will ask what kind of questions, and also interest level of the audience. Based on these analyses, we design a smart posterboard which can sense human behaviors and annotate interactions and interest level during poster sessions.

An End-to-End Evaluation of Two Situated Dialog Systems
Lina M. Rojas Barahona, Alejandra Lorenzo and Claire Gardent
Thursday July 5, 2012 10:45–11:10

We present and evaluate two state-of-the art dialogue systems developed to support dialog with French speaking virtual characters in the context of a serious game: one hybrid statistical/symbolic and one purely statistical. In evaluating the two systems, we seek to compare their usability. We also seek to better understand which module causes which errors and with which impact on usability. To address these questions, we conducted a quantitative evaluation where we compare the accuracy of the interpreter and of the dialog manager used by each system; a user based evaluation based on 22 subjects using both the QA and the hybrid system; and a corpus based evaluation where we examine such criteria as dialog coherence, dialog success, interpretation and generation errors in the corpus of Human-System interactions collected during the user-based evaluation. We show that although the statistical approach is slightly more robust, the hybrid strategy seems to be better at guiding the player through the game.

"Love ya, jerkface": Using Sparse Log-Linear Models to Build Positive (and Impolite) Relationships with Teens
William Yang Wang, Samantha Finkelstein, Amy Ogan, Alan W. Black and Justine Cassell
Thursday July 5, 2012 11:10–11:35

One challenge of implementing spoken dialogue systems for long-term interaction is how to adapt the dialogue as user and system become more familiar. We believe this challenge includes evoking and signaling aspects of long-term relationships such as rapport. For tutoring systems, this may additionally require knowing how relationships are signaled among non-adult users. We therefore investigate conversational strategies used by teenagers in peer tutoring dialogues, and how these strategies function differently among friends or strangers. In particular, we use annotated and automatically extracted linguistic devices to predict impoliteness and positivity in the next turn. To take into account the sparse nature of these features in real data we use models including Lasso, ridge estimator, and Elastic Net. We evaluate the predictive power of our models under various settings, and compare our sparse models with standard non-sparse solutions. Our experiments demonstrate that our models are more accurate than non-sparse models quantitatively, and that teens use unexpected kinds of language to do relationship work such as signaling rapport, but friends and strangers, tutors and tutees, carry out this work in quite different ways from one another.

Enhancing Referential Success by Tracking Hearer Gaze

Alexander Koller, Konstantina Garoufi, Maria Staudte and Matthew Crocker

Thursday July 5, 2012 11:35–12:00

The ability to monitor the communicative success of its utterances and, if necessary, provide feedback and repair is useful for a dialog system. We show that in situated communication, eyetracking can be used to reliably and efficiently monitor the hearer's reference resolution process. An interactive system that draws on hearer gaze to provide positive or negative feedback after referring to objects outperforms baseline systems on metrics of referential success and user confusion.

Unsupervised Topic Modeling Approaches to Decision Summarization in Spoken Meetings

Lu Wang and Claire Cardie

Thursday July 5, 2012 1:00–1:25

We present a token-level decision summarization framework which utilizes the latent topic structures of utterances to identify the summary-worthy words. Concretely, a series of unsupervised topic models are explored and experimental results show that fine-grained topic models which discover topics on utterance-level can better identify the gist information in the decision-making process. Moreover, our proposed token-level summarization approaches, which is able to remove redundancies within utterances, outperforms existing utterance ranking based summarization methods. finally, context information is also investigated to add more relevant information for summary.

An Unsupervised Approach to User Simulation: Toward Self-Improving Dialog Systems

Sungjin Lee and Maxine Eskenazi

Thursday July 5, 2012 1:25–1:50

This paper proposes an unsupervised approach to user simulation in order to automatically furnish updates and assessments of a deployed spoken dialog system. The proposed method adopts a dynamic Bayesian network to infer the unobservable true user action from which the parameters of other components are naturally derived. To verify the quality of the simulation, the proposed method was applied to the Let's Go domain and a set of measures was used to analyze the simulated data at several levels. The results showed a very close correspondence between the real and simulated data, implying that it is possible to create a realistic user simulator that does not necessitate human intervention.

Hierarchical Conversation Structure Prediction in Multi-Party Chat

Elijah Mayfield, David Adamson and Carolyn Penstein Rosé

Thursday July 5, 2012 1:50–2:15

Conversational practices do not occur at a single unit of analysis. To understand the interplay between social positioning, information sharing, and rhetorical strategy in language, various granularities are necessary. In this work we present a machine learning model for multi-party chat which predicts conversation structure across differing units of analysis. first, we mark sentence-level behavior using an information sharing annotation scheme. By taking advantage of Integer Linear Programming and a sociolinguistic framework, we enforce structural relationships between sentence-level annotations and sequences of interaction. Then, we show that clustering these sequences can effectively disentangle the threads of conversation. This model is highly accurate, performing near human accuracy, and performs analysis on-line, opening the door to real-time analysis of the discourse of conversation.

Rapid Development Process of Spoken Dialogue Systems using Collaboratively Constructed Semantic Resources
Masahiro Araki
Thursday July 5, 2012 2:30–4:00

We herein propose a method for the rapid development of a spoken dialogue system based on collaboratively constructed semantic resources and compare the proposed method with a conventional method that is based on a relational database. Previous development frameworks of spoken dialogue systems, which presuppose a relational database management system as a background application, require complex data definition, such as making entries in a task-dependent language dictionary, templates of semantic frames, and conversion rules from user utterances to the query language of the database. We demonstrate that a semantic web oriented approach based on collaboratively constructed semantic resources significantly reduces troublesome rule descriptions and complex configurations in the rapid development process of spoken dialogue systems.

The Effect of Cognitive Load on a Statistical Dialogue System
Milica Gašić, Pirros Tsiakoulis, Matthew Henderson, Blaise Thomson, Kai Yu, Eli Tzirkel and Steve Young
Thursday July 5, 2012 2:30–4:00

In recent years statistical dialogue systems have gained significant attention due to their potential to be more robust to speech recognition errors. However, for automotive and similar hands-free applications, these systems must also be robust to changes in user behaviour caused by cognitive loading. In this paper, a statistical dialogue system providing restaurant information is evaluated in a set-up where the subjects used a driving simulator whilst talking to the system. The influences of cognitive loading were investigated in some detail and some clear differences in behaviour were discovered. In particular, it was found that users chose to respond to different system questions and use different speaking styles. These behavioural changes indicate the need for on-line policy adaptation and an incremental dialogue approach rather than the simple turn-taking model used in the current system.

Predicting Adherence to Treatment for Schizophrenia from Dialogue Transcripts
Christine Howes, Matthew Purver, Rose McCabe, Patrick G. T. Healey and Mary Lavelle
Thursday July 5, 2012 2:30–4:00

Recent work on consultations between outpatients with schizophrenia and psychiatrists has shown that adherence to treatment can be predicted by patterns of repair – specifically, the pro-activity of the patient in checking their understanding, i.e. patient clarification. Using machine learning techniques, we investigate whether this tendency can be predicted from high-level dialogue features, such as backchannels, overlap and the proportion of talk of each participant. The results indicate that these features are not predictive of patient adherence to treatment or patient's satisfaction with the communication, although they do have some association with patient symptoms. However, all these can be predicted if we allow features at the individual word level. These preliminary experiments indicate that patient adherence is predictable from dialogue transcripts, but further work is necessary to develop a meaningful, general and reliable feature set.

Reinforcement Learning of Question-Answering Dialogue Policies for Virtual Museum Guides

Teruhisa Misu, Kallirroi Georgila, Anton Leuski and David Traum
Thursday July 5, 2012 2:30–4:00

We use Reinforcement Learning (RL) to learn question-answering dialogue policies for a real-world application. We analyze a corpus of interactions of museum visitors with two virtual characters that serve as guides at the Museum of Science in Boston, in order to build a realistic model of user behavior when interacting with these characters. A simulated user is built based on this model and used for learning the dialogue policy of the virtual characters using RL. Our learned policy outperforms two baselines (including the original dialogue policy that was used for collecting the corpus) in a simulation setting.

From Strangers to Partners: Examining Convergence within a Longitudinal Study of Task-Oriented Dialogue

Christopher M. Mitchell, Kristy Elizabeth Boyer and James C. Lester
Thursday July 5, 2012 2:30–4:00

Convergence is thought to be an important phenomenon in dialogue through which interlocutors adapt to each other. Yet, its mechanisms and relationship to dialogue outcomes are not fully understood. This paper explores convergence in textual task-oriented dialogue during a longitudinal study. The results suggest that over time, convergence between interlocutors increases with successive dialogues. Additionally, for the tutorial dialogue domain at hand, convergence metrics were found to be significant predictors of dialogue outcomes such as learning, mental effort, and emotional states including frustration, boredom, and confusion. The results suggest ways in which dialogue systems may leverage convergence to enhance their interactions with users.

The Structure and Generality of Spoken Route Instructions

Aasish Pappu and Alexander Rudnicky
Thursday July 5, 2012 2:30-4:00

A robust system that understands route instructions should be able to process instructions generated naturally by humans. A desirable property would be for it to handle repairs and other modifications to existing instructions. To this end, we collected a corpus of spoken instructions (and modified instructions) produced by subjects provided with an origin and a destination. We found that instructions could be classified into four categories, depending on their intent such as imperative, feedback, or meta comment. We asked a different set of subjects to follow these directions to determine the usefulness and comprehensibility of individual instructions. finally, we used a part of these data to construct a semantic grammar and evaluated its coverage on the collected corpus. To determine whether instruction giving forms a predictable sub-language, we tested the grammar on three corpora collected by others and determined that this was largely the case. Our work suggests that predictable sub-language may exist for well-defined tasks.

Improving Implicit Discourse Relation Recognition Through Feature Set Optimization

Joonsuk Park and Claire Cardie
Thursday July 5, 2012 2:30–4:00

We provide a systematic study of previously proposed features for implicit discourse relation identification, identifying new feature combinations that optimize F1-score. The resulting classifiers achieve the best F1-scores to date for the four top-level discourse relation classes of the

Penn Discourse Tree Bank: COMPARISON, CONTINGENCY, EXPANSION, and TEMPORAL. We further identify factors for feature extraction that can have a major impact on the performance and determine that some features originally proposed for the task no longer provide performance gains in light of more powerful, recently discovered features. Our results constitute a new set of baselines for future studies of implicit discourse relation identification.

A Temporal Simulator for Developing Turn-Taking Methods for Spoken Dialogue Systems
Ethan O. Selfridge and Peter A. Heeman
Thursday July 5, 2012 2:30–4:00

Developing sophisticated turn-taking behavior is necessary for next-generation dialogue systems. However, incorporating real users into the development cycle is expensive and current simulation techniques are inadequate. As a foundation for advancing turn-taking behavior we present a temporal simulator that models user speech, natural language generation, incremental speech recognition, and voice activity detection. We describe the details of the simulator and demonstrate it on a sample domain.

Dialogue Act Recognition using Reweighted Speaker Adaptation
Congkai Sun and Louis-Philippe Morency
Thursday July 5, 2012 2:30–4:00

In this work we study the effectiveness of speaker adaptation for dialogue act recognition in multiparty meetings. first, we analyze idiosyncracy in dialogue verbal acts by qualitatively studying the differences and conflicts among speakers and by quantitively comparing speaker-specific models. Based on these observations, we propose a new approach for dialogue act recognition based on reweighted domain adaptation which effectively balance the influence of speaker specific and other speakers' data. Our experiments on a real-world meeting dataset show that with even only 200 speaker-specific annotated dialogue acts, the performances on dialogue act recognition are significantly improved when compared to several baseline algorithms. To our knowledge, this work is the first to tackle this promising research direction of speaker adaptation for dialogue act recognition.

Estimating Adaptation of Dialogue Partners with Different Verbal Intelligence
Kseniya Zablotskaya, Fernando Fernández-Martínez and Wolfgang Minker
Thursday July 5, 2012 2:30–4:00

This work investigates to what degree speakers with different verbal intelligence may adapt to each other. The work is based on a corpus consisting of 100 descriptions of a short film (monologues), 56 discussions about the same topic (dialogues), and verbal intelligence scores of the test participants. Adaptation between two dialogue partners was measured using cross-referencing, proportion of "I", "You" and "We" words, between-subject correlation and similarity of texts. It was shown that lower verbal intelligence speakers repeated more nouns and adjectives from the other and used the same linguistic categories more often than higher verbal intelligence speakers. In dialogues between strangers, participants with higher verbal intelligence showed a greater level of adaptation.

A Demonstration of Incremental Speech Understanding and Confidence Estimation in a Virtual Human Dialogue System

David DeVault and David Traum
Thursday July 5, 2012 2:30–4:00

This demonstration highlights some emerging capabilities for incremental speech understanding and processing in virtual human dialogue systems. Our approach to incremental speech understanding enables virtual humans to predict the meaning of user utterances while they are still in progress, and to assign confidence to these predictions. We show how we have applied these techniques in a new virtual human negotiation domain, and discuss some possible applications.

Integrating Location, Visibility, and Question-Answering in a Spoken Dialogue System for Pedestrian City Exploration

Srinivasan Janarthanam, Oliver Lemon, Xingkun Liu, Phil Bartie, William Mackaness,
Tiphaine Dalmas and Jana Goetze
Thursday July 5, 2012 2:30–4:00

We demonstrate a spoken dialogue-based information system for pedestrians. The system is novel in combining geographic information system (GIS) modules such as a Visibility Engine with a Question-Answering (QA) system, integrated within a dialogue system architecture. Users of the demonstration system can use a web-based version (simulating pedestrian movement using StreetView) to engage in a variety of interleaved conversations such as navigating from A to B, using the QA functionality to learn more about points of interest (PoI) nearby, and searching for amenities and tourist attractions. This system explores a variety of research questions involving the integration of multiple information sources within conversational interaction.

A Mixed-Initiative Conversational Dialogue System for Healthcare

Fabrizio Morbini, Eric Forbell, David DeVault, Kenji Sagae, David Traum and Albert Rizzo
Thursday July 5, 2012 2:30–4:00

We present a mixed initiative conversational dialogue system designed to address primarily mental health care concerns related to military deployment. It is supported by a new information-state based dialogue manager, FLoReS (Forward-Looking, Reward Seeking dialogue manager), that allows both advanced, flexible, mixed initiative interaction, and efficient policy creation by domain experts. To easily reach its target population this dialogue system is accessible as a web application.

Towards Mediating Shared Perceptual Basis in Situated Dialogue
Changsong Liu, Rui Fang and Joyce Chai
Thursday July 5, 2012 4:25–4:50

To enable effective referential grounding in situated human robot dialogue, we have conducted an empirical study to investigate how conversation partners collaborate and mediate shared basis when they have mismatched visual perceptual capabilities. In particular, we have developed a graph-based representation to capture linguistic discourse and visual discourse, and applied inexact graph matching to ground references. Our empirical results have shown that, even when computer vision algorithms produce many errors (e.g. 84:7% of the objects in the environment are mis-recognized), our approach can still achieve 66% accuracy in referential grounding. These results demonstrate that, due to its error-tolerance nature, inexact graph matching provides a potential solution to mediate shared perceptual basis for referential grounding in situated interaction.

Global Features for Shallow Discourse Parsing
Sucheta Ghosh, Giuseppe Riccardi and Richard Johansson
Thursday July 5, 2012 4:50–5:15

A coherently related group of sentences may be referred to as a discourse. In this paper we address the problem of parsing coherence relations as defined in the Penn Discourse Tree Bank (PDTB). A good model for discourse structure analysis needs to account both for local dependencies at the token-level and for global dependencies and statistics. We present techniques on using inter-sentential or sentence-level (global), data-driven, non-grammatical features in the task of parsing discourse. The parser model follows up previous approach based on using token-level (local) features with conditional random fields for shallow discourse parsing, which is lacking in structural knowledge of discourse. The parser adopts a two-stage approach where first the local constraints are applied and then global constraints are used on a reduced weighted search space (n-best). In the latter stage we experiment with different rerankers trained on the first stage n-best parses, which are generated using lexico-syntactic local features. The two-stage parser yields significant improvements over the best performing model of discourse parser on the PDTB corpus.

A Reranking Model for Discourse Segmentation using Subtree Features
Ngo Xuan Bach, Nguyen Le Minh and Akira Shimazu
Thursday July 5, 2012 5:15–5:40

This paper presents a discriminative reranking model for the discourse segmentation task, the first step in a discourse parsing system. Our model exploits subtree features to rerank N-best outputs of a base segmenter, which uses syntactic and lexical features in a CRF framework. Experimental results on the RST Discourse Treebank corpus show that our model outperforms existing discourse segmenters in both settings that use gold standard Penn Treebank parse trees and Stanford parse trees.

Landmark-Based Location Belief Tracking in a Spoken Dialog System

Yi Ma, Antoine Raux, Deepak Ramachandran and Rakesh Gupta
Friday July 6, 2012 9:15–9:40

Many modern spoken dialog systems use probabilistic graphical models to update their belief over the concepts under discussion, increasing robustness in the face of noisy input. However, such models are ill-suited to probabilistic reasoning about spatial relationships between entities. In particular, a car navigation system that infers users' intended destination using nearby landmarks as descriptions must be able to use distance measures as a factor in inference. In this paper, we describe a belief tracking system for a location identification task that combines a semantic belief tracker for categorical concepts based on the DPOT framework (Raux and Ma, 2011) with a kernel density estimator that incorporates landmark evidence from multiple turns and landmark hypotheses, into a posterior probability over candidate locations. We evaluate our approach on a corpus of destination setting dialogs and show that it significantly outperforms a deterministic baseline.

Probabilistic Dialogue Models with Prior Domain Knowledge

Pierre Lison
Friday July 6, 2012 9:40–10:05

Probabilistic models such as Bayesian Networks are now in widespread use in spoken dialogue systems, but their scalability to complex interaction domains remains a challenge. One central limitation is that the state space of such models grows exponentially with the problem size, which makes parameter estimation increasingly difficult, especially for domains where only limited training data is available. In this paper, we show how to capture the underlying structure of a dialogue domain in terms of probabilistic rules operating on the dialogue state. The probabilistic rules are associated with a small, compact set of parameters that can be directly estimated from data. We argue that the introduction of this abstraction mechanism yields probabilistic models that are easier to learn and generalise better than their unstructured counterparts. We empirically demonstrate the benefits of such an approach learning a dialogue policy for a human-robot interaction domain based on a Wizard-of-Oz data set.

Exploiting Machine-Transcribed Dialog Corpus to Improve Multiple Dialog States Tracking Methods

Sungjin Lee and Maxine Eskenazi
Friday July 6, 2012 10:05–10:30

This paper proposes the use of unsupervised approaches to improve components of partition-based belief tracking systems. The proposed method adopts a dynamic Bayesian network to learn the user action model directly from a machine-transcribed dialog corpus. It also addresses confidence score calibration to improve the observation model in a unsupervised manner using dialog-level grounding information. To verify the effectiveness of the proposed method, we applied it to the Let's Go domain. Overall system performance for several comparative models were measured. The results show that the proposed method can learn an effective user action model without human intervention. In addition, the calibrated confidence score was verified by demonstrating the positive influence on the user action model learning process and on overall system performance.

Cohesion, Entrainment and Task Success in Educational Dialog
Diane Litman
Friday July 6, 2012 11:00–12:00

Researchers often study dialog corpora to better understand what makes some dialogs more successful than others. In this talk I will examine the relationship between coherence/entrainment and task success, in several types of educational dialog corpora: 1) one-on-one tutoring, where students use dialog to interact with a human tutor in the physics domain, 2) one-on-one tutoring, where students instead interact with a spoken dialog system, and 3) engineering design, where student teams engage in multi-party dialog to complete a group project. I will first introduce several corpus-based measures of both lexical and acoustic-prosodic dialog cohesion and entrainment, and extend them to handle multi-party conversations. I will then show that the amount of cohesion and/or entrainment positively correlates with measures of educational task success in all of our corpora. finally, I will discuss how we are using our findings to build better tutorial dialog systems.

A Bottom-Up Exploration of the Dimensions of Dialog State in Spoken Interaction
Nigel G. Ward and Alejandro Vega
Friday July 6, 2012 1:30–1:55

Models of dialog state are important, both scientifically and practically, but today's best build strongly on tradition. This paper presents a new way to identify the important dimensions of dialog state, more bottom-up and empirical than previous approaches. Specifically, we applied Principal Component Analysis to a large number of low-level prosodic features to find the most important dimensions of variation. The top 20 out of 76 dimensions accounted for 81% of the variance, and each of these dimensions clearly related to dialog states and activities, including turn taking, topic structure, grounding, empathy, cognitive processes, attitude and rhetorical structure.

Using Group History to Identify Character-Directed Utterances in Multi-Child Interactions
Hannaneh Hajishirzi, Jill F. Lehman and Jessica K. Hodgins
Friday July 6, 2012 1:55–2:20

Addressee identification is an element of all language-based interactions, and is critical for turn-taking. We examine the particular problem of identifying when each child playing an interactive game in a small group is speaking to an animated character. After analyzing child and adult behavior, we explore a family of machine learning models to integrate audio and visual features with temporal group interactions and limited, task-independent language. The best model performs identification about 20% better than the model that uses the audio-visual features of the child alone.

Adapting to Multiple Affective States in Spoken Dialogue
Kate Forbes-Riley and Diane Litman
Friday July 6, 2012 2:20–2:45

We evaluate a wizard-of-oz spoken dialogue system that adapts to multiple user affective states in real-time: user disengagement and uncertainty. We compare this version with the prior version of our system, which only adapts to user uncertainty. Our analysis investigates how iteratively adding new affect adaptation to an existing affect-adaptive system impacts global and local performance. We find a significant increase in motivation for users who most frequently

received the disengagement adaptation. Moreover, responding to disengagement breaks its negative correlations with task success and user satisfaction, reduces uncertainty levels, and reduces the likelihood of continued disengagement.

Dialog System Using Real-Time Crowdsourcing and Twitter Large-Scale Corpus
Fumihiro Bessho, Tatsuya Harada and Yasuo Kuniyoshi
Friday July 6, 2012 3:00–4:00

We propose a dialog system that creates responses based on a large-scale dialog corpus retrieved from Twitter and real-time crowdsourcing. Instead of using complex dialog management, our system replies with the utterance from the database that is most similar to the user input. We also propose a real-time crowdsourcing framework for handling the case in which there is no adequate response in the database.

Automatically Acquiring Fine-Grained Information Status Distinctions in German
Aoife Cahill and Arndt Riester
Friday July 6, 2012 3:00–4:00

We present a model for automatically predicting information status labels for German referring expressions. We train a CRF on manually annotated phrases, and predict a fine-grained set of labels. We achieve an accuracy score of 69.56% on our most detailed label set, 76.62% when gold standard coreference is available.

A Unified Probabilistic Approach to Referring Expressions
Kotaro Funakoshi, Mikio Nakano, Takenobu Tokunaga and Ryu Iida
Friday July 6, 2012 3:00–4:00

This paper proposes a probabilistic approach to the resolution of referring expressions for task-oriented dialogue systems. The approach resolves descriptions, anaphora, and deixis in a unified manner. In this approach, the notion of reference domains serves an important role to handle context-dependent attributes of entities and references to sets. The evaluation with the REX-J corpus shows promising results.

Combining Verbal and Nonverbal Features to Overcome the "Information Gap" in Task-Oriented Dialogue
Eun Young Ha, Joseph F. Grafsgaard, Christopher M. Mitchell, Kristy Elizabeth Boyer and James C. Lester
Friday July 6, 2012 3:00–4:00

Dialogue act modeling in task-oriented dialogue poses significant challenges. It is particularly challenging for corpora consisting of two interleaved communication streams: a dialogue stream and a task stream. In such corpora, information can be conveyed implicitly by the task stream, yielding a dialogue stream with seemingly missing information. A promising approach leverages rich resources from both the dialog and the task streams, combining verbal and non-verbal features. This paper presents work on dialogue act modeling that leverages body posture, which may be indicative of particular dialogue acts. Combining three information sources (dialogue exchanges, task context, and users' posture), three types of machine learning frameworks were compared. The results indicate that some models better preserve the structure of task-oriented dialogue than others, and that automatically recognized postural features may help to disambiguate user dialogue moves.

Semantic Specificity in Spoken Dialogue Requests

Ben Hixon, Rebecca J. Passonneau and Susan L. Epstein
Friday July 6, 2012 3:00–4:00

Ambiguous or open-ended requests to a dialogue system result in more complex dialogues. We present a semantic-specificity metric to gauge this complexity for dialogue systems that access a relational database. An experiment where a simulated user makes requests to a dialogue system shows that semantic specificity correlates with dialogue length.

Contingency and Comparison Relation Labeling and Structure Prediction in Chinese Sentences

Hen-Hsen Huang and Hsin-Hsi Chen
Friday July 6, 2012 3:00–4:00

Unlike in English, the sentence boundaries in Chinese are fuzzy and not well-defined. As a result, Chinese sentences tend to be long and consist of complex discourse relations. In this paper, we focus on two important relations, Contingency and Comparison, which occur often inside a sentence. We construct a moderate-sized corpus for the investigation of intra-sentential relations and propose models to label the relation structure. A learning based model is evaluated with various features. Experimental results show our model achieves accuracies of 81.63% in the task of relation labeling and 74.8% in the task of relation structure prediction.

A Study in How NLU Performance Can Affect the Choice of Dialogue System Architecture

Anton Leuski and David DeVault
Friday July 6, 2012 3:00–4:00

This paper presents an analysis of how the level of performance achievable by an NLU module can affect the optimal modular design of a dialogue system. We present an evaluation that shows how NLU accuracy levels impact the overall performance of a system that includes an NLU module and a rule-based dialogue policy. We contrast these performance levels with the performance of a direct classification design that omits a separate NLU module. We conclude with a discussion of the potential for a hybrid architecture incorporating the strengths of both approaches.

Integrating Incremental Speech Recognition and POMDP-Based Dialogue Systems

Ethan O. Selfridge, Iker Arizmendi, Peter A. Heeman and Jason D. Williams
Friday July 6, 2012 3:00–4:00

The goal of this paper is to present a first step toward integrating Incremental Speech Recognition (ISR) and Partially-Observable Markov Decision Process (POMDP) based dialogue systems. The former provides support for advanced turn-taking behavior while the other increases the semantic accuracy of speech recognition results. We present an Incremental Interaction Manager that supports the use of ISR with strictly turn-based dialogue managers. We then show that using a POMDP-based dialogue manager with ISR substantially improves the semantic accuracy of the incremental results.

A Regression-Based Approach to Modeling Addressee Backchannels
Allison Terrell and Bilge Mutlu
Friday July 6, 2012 3:00–4:00

During conversations, addressees produce conversational acts – verbal and nonverbal backchannels – that facilitate turn-taking, acknowledge speakership, and communicate common ground without disrupting the speaker's speech. These acts play a key role in achieving fluent conversations. Therefore, gaining a deeper understanding of how these acts interact with speaker behaviors in shaping conversations might offer key insights into the design of technologies such as computer-mediated communication systems and embodied conversational agents. In this paper, we explore how a regression-based approach might offer such insights into modeling predictive relationships between speaker behaviors and addressee backchannels in a storytelling scenario. Our results reveal speaker eye contact as a significant predictor of verbal, nonverbal, and bimodal backchannels and utterance boundaries as predictors of nonverbal and bimodal backchannels.

Improving Sentence Completion in Dialogues with Multi-Modal Features
Anruo Wang, Barbara Di Eugenio and Lin Chen
Friday July 6, 2012 3:00–4:00

With the aim of investigating how humans understand each other through language and gestures, this paper focuses on how people understand incomplete sentences. We trained a system based on interrupted but resumed sentences, in order to find plausible completions for incomplete sentences. Our promising results are based on multi-modal features.

Combining Incremental Language Generation and Incremental Speech Synthesis for Adaptive Information Presentation
Hendrik Buschmeier, Timo Baumann, Benjamin Dosch, Stefan Kopp and David Schlangen
Friday July 6, 2012 4:25–4:50

Participants in a conversation are normally receptive to their surroundings and their interlocutors, even while they are speaking and can, if necessary, adapt their ongoing utterance. Typical dialogue systems are not receptive and cannot adapt while uttering. We present combinable components for incremental natural language generation and incremental speech synthesis and demonstrate the flexibility they can achieve with an example system that adapts to a listener's acoustic understanding problems by pausing, repeating and possibly rephrasing problematic parts of an utterance. In an evaluation, this system was rated as significantly more natural than two systems representing the current state of the art that either ignore the interrupting event or just pause; it also has a lower response time.

Focused Meeting Summarization via Unsupervised Relation Extraction
Lu Wang and Claire Cardie
Friday July 6, 2012 4:50–5:15

We present a novel unsupervised framework for focused meeting summarization that views the problem as an instance of relation extraction. We adapt an existing in-domain relation learner (Chen et al., 2011) by exploiting a set of task-specific constraints and features. We evaluate the approach on a decision summarization task and show that it outperforms unsupervised utterance-level extractive summarization baselines as well as an existing generic relation-extraction-based summarization method. Moreover, our approach produces summaries competitive with those generated by supervised methods in terms of the standard ROUGE score.

Markov Logic Networks for Situated Incremental Natural Language Understanding

Casey Kennington and David Schlangen

Friday July 6, 2012 5:15–5:40

We present work on understanding natural language in a situated domain, that is, language that possibly refers to visually present entities, in an incremental, word-by-word fashion. Such type of understanding is required in conversational systems that need to act immediately on language input, such as multi-modal systems or dialogue systems for robots. We explore a set of models specified as Markov Logic Networks, and show that a model that has access to information about the visual context of an utterance, its discourse context, as well as the linguistic structure of the utterance performs best. We explore its incremental properties, and also its use in a joint parsing and understanding module. We conclude that MLNs offer a promising framework for specifying such models in a general, possibly domain-independent way.

Index

In cooperation with:

Association for Computational Linguistics (ACL)

International Speech Communication Association (ISCA)

Association for the Advancement of Artificial Intelligence (AAAI)

Sponsors:

AT&T

AVIOS

Honda Research Institute

IBM Research

KT Corporation

Microsoft Research

NHN Corporation

Seoul National University

http://www.sigdial.org/workshops/conference13