

Multi-modal dialogue systems for industrial applications

Arne Jönsson
Institutionen för datavetenskap

Sammanfattning av de vetenskapliga resultaten

Forskningen har bedrivits inom flera olika områden som alla på olika sätt relaterar till multi-modal dialog.

- Vi har byggt ett antal multi-modala dialogsystem samt studerat användningen av dessa i olika situationer [41, 1, 2, 3, 3, 33]. Mycket av detta gjordes i samarbete med Nokia Home Communications [23, 22, 24, 27].
Ett resultat värt att peka ut här är att vi visat att användares uppfattning av ett system beror mycket på om det uppfattas som verktyg eller kommunikationspartner [40, 37, 39].
- Vi har, tillsammans med IBM Almaden i USA, utvecklat världens första helt ögonrörelsestyrda dialogsystem. Systemet utvärderades och resultaten visar att det fungerar och att användarna är positiva till systemet [36, 42, 39]. Systemet utvecklades baserat på studier av vilken funktion ögonrörelser har i vanlig interaktion.
- Vi har utvecklat nya tekniker för att förbättra taligenkänning genom att kombinera statistiskt baserade språkmodeller med grammatikbaserade [18, 17, 19, 20]. Denna teknik har använts för att identifiera okända ord, för att ge riktad hjälp samt för att ge förslag på korrekta sätt att uttrycka sig.
- I arbetet med dialogsystem utvecklade vi en ny arkitektur som bland annat separerar domänkunskaps hantering från dialoghantering [10, 11, 14, 9, 13, 15]. Detta gör att vi i dialogsystemet kan integrera information från flera datakällor utan att det behöver påverka dialoghanteringen.
- Den nya arkitekturen gjorde det också enklare att använda ontologier och vi utvecklade ett formellt ramverk för användning av ontologier i dialogsystem [12, 8].
- Tillsammans med Sveriges Television utvecklade vi ett dialogsystem där databasen extraherats automatiskt ur ett textdokument, i det här fallet en fågelbok [31, 30, 16]. Detta var första gången någon provat att kombinera de två olika språkteknologiska forskningsområdena dokumentanalys och interaktionssystemutveckling. Detta föregicks av arbete med Ida Infront kring utveckling av informationssystem för myndighetstexter [34, 35].

- Vi utvecklade en modul för att integrera talad och penn-inmatad information [25, 32].
- För att underlätta användningen av våra resultat i industrin, skapade vi den öppna källkodsresursen NLPFARM [7] och utvecklade metoder för att evolutionärt bygga dialogsystem [29, 5, 6]. Som en del av detta utvecklade vi PGP (Phase Graph Process) som är ett designmönster för att enkelt skapa dialogsystem genom att kombinera processmoduler i en fasgraf [4]. NLPFARM har haft drygt 200000 besökare och 4500 nedladdningar (oktober 2005).
- Vi har studerat adaptivitet och hur man kan använda en modell av användaren för att göra interaktionen effektivare. Detta gjordes i ett adaptivt multimodalt rekommendationssystem som i sin första version hanterar kallstartsproblemet genom multimodal dialog [27, 26, 28, 43, 45]. Metoden har också utvärderats [44].

Examina och befordringar som projektet bidragit till

Inom projektet har två doktorsavhandlingar presenterats, Annika Flycht-Eriksson [8] och Pernilla Qvarfordt [39] och fyra licavhandlingar [11, 38, 28, 19].

Projektledaren har befordrats till professor.

Examensarbeten som genomförts inom ramen för projektet

Syed M Zaidi, Design and Development of Interactive 2D Multimodal Dialogue Partner, 2005.

Anna Ekström, Automatisk klassificering av nyhetstelegram, en jämförelse av två tillvägagångssätt, 2004.

Anna Hjalmarsson, Evaluating AdApt, a multi-modal conversational, dialogue system, using PARADISE, 2003.

Jenny Isberg och Sophie Öhrn, Designprocess och dialoginsamling för adaptiva dialogsystem, 2003.

Petronella Andersson, User's reactions to interactive agents when their visual and verbal expressions do not match, 2002.

Jonas Kvist, Dialoghantering och e-handel med naturligt språk, 2001.

Håkan Johansson, Integrating Multimodal Information, 2001.

Pontus Johansson, Iterative Development of an Information-Providing Dialogue System, 2001.

Garbirel Skantze, Koordinering av refererande uttryck i multimodal människa-datordialog, 2000.

Martin Gustavsson, Designing a multimodal system for a culturally diverse user group, 2000.

Personer som finansierats av projektet

Genom åren har, utöver projektledaren, olika personer finansierats inom projektet. 2000 var det Nils Dahlbäck, Annika Flycht-Eriksson och Pernilla Qvarfordt. 2001 finansierades inte längre Nils Dahlbäck inom projektet istället tillkom Lars Degerstedt och Håkan Johansson som var den första doktoranden inom projektet som också var finansierad inom den nationella forskarskolan i språkteknologi (GSLT). Från 2002 delfinansierades GSLT-doktoranderna Pontus Johansson och Genevieve Gorrell istället för Håkan Johansson. Samtidigt minskade finansieringen av Flycht-Eriksson och Qvarfordt.

Industrikontakter

Från start har vi samarbetat med Telia research och gruppen kopplad till deras språkteknologilaboratorium.

2001 etablerades samarbete med Nokia Home Communications. Detta pågick fram till dess att Nokia lade ner verksamheten i Mjärdevi. Inom detta samarbetet gjorde vi studier av multimodal medieterminals-interaktion. Vi byggde också en interaktiv TV-guide.

Vi försökte etablera ett samarbete med Kreatel AB och Zenterio AB som också arbetade med medieterminaler, men detta ledde inte till något konkret projekt.

Vi har under hela projekttiden haft samarbete med Ida Infront AB. Under 2005 etablerades också samarbete med Humany AB och vi har två forskningsanslagsansökningar tillsammans.

Övrigt

Det formades en forskargrupp som budgetmässigt sorterades under Laboratoriet för databehandling av naturligt språk men som verksamhetsmässigt var en egen enhet. Forskargruppen har under perioden finansierats av olika externa anslag från Svenska IT Institutet, Vinnova, KFB och HSFR/NUTEK om totalt 15,5 miljoner kronor utöver CENIIT-projektet. Samtliga dessa projekt har på olika sätt handlat om multimodal interaktion. Som mest bestod forskargruppen av Lars Degerstedt som är disputerad i logikprogrammering vid LiTH och nu är 1:e forskningsingenjör samt doktoranderna Annika Flycht-Eriksson, Pernilla Qvarfordt, Pontus Johansson, Håkan Johansson och Genevieve Gorrell. Magnus Merkel och Nils Dahlbäck (båda disputerade forskare) har också varit knutna till gruppen under vissa perioder.

Forskargruppen består för närvarande av Lars Degerstedt, Pontus Wärnestål (fd. Johansson), Annika Silvervarg (fd. Flycht-Eriksson) samt Genevieve Gorrell.

Forskargruppen hade under en period kontakter med Ankica Babics Ceniit-projekt, vilket resulterade i en publikation [21].

References

- [1] Patrik Bäckvall, Per Mårtensson, and Pernilla Qvarfordt. Using fisheye for navigation on small displays. In *Proceedings of NordiCHI 2000, Stockholm, Sweden, 2000*.
- [2] Aseel Berglund and Pernilla Qvarfordt. Error resolution strategies for interactive television speech interfaces. In *Proceedings of INTERACT 2003, Zürich, Switzerland, 2003*.
- [3] Nils Dahlbäck and Arne Jönsson. Experiences with and lessons learned from working with a modular natural language dialogue architecture. In *Proceedings of HCI International 2003, 2003*.
- [4] Lars Degerstedt and Pontus Johansson. Evolutionary Development of Phase-Based Dialogue Systems. In *Proc. of the 8th Scandinavian Conference on Artificial Intelligence*, pages 59–67, Bergen, Norway, November 2003.
- [5] Lars Degerstedt and Arne Jönsson. A Method for Systematic Implementation of Dialogue Management. In *Workshop notes from the 2nd IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems, Seattle, WA, 2001*.
- [6] Lars Degerstedt and Arne Jönsson. Iterative Implementation of Dialogue System Modules. In *Proceedings of Eurospeech 2001, Aalborg, Denmark, 2001*.
- [7] Lars Degerstedt and Arne Jönsson. Open Resources for Language Technology. In *Proceedings of 4th International Conference on Language Resources and Evaluation, LREC 2004, Lisboa, Portugal, 2004*.
- [8] Annika Flycht-Ericsson. *Design and use of Ontologies in Information-providing Dialogue Systems*. PhD thesis, Linköping University, 2004.
- [9] Annika Flycht-Eriksson. A survey of knowledge sources in dialogue systems. In *Proceedings of IJCAI-99 Workshop on Knowledge and Reasoning in Practical Dialogue Systems, August, Stockholm, 1999*.
- [10] Annika Flycht-Eriksson. A domain knowledge manager for dialogue systems. In *Proceedings of the 14th European Conference on Artificial Intelligence, ECAI 2000*. IOS Press, Amsterdam, 2000.
- [11] Annika Flycht-Eriksson. Domain knowledge management in information-providing dialogue systems. Licentiate Thesis 890, Linköping Studies in Science and Technology, Linköping University, 2001.
- [12] Annika Flycht-Eriksson. Design of ontologies for dialogue interaction and information extraction. In *Proceedings of IJCAI'03 workshop on Knowledge and reasoning in practical dialogue systems, Acapulco, Mexico, 2003*.
- [13] Annika Flycht-Eriksson. Representing knowledge of dialogue, domain, task and user in dialogue systems - how and why? *Electronic Transactions on Artificial Intelligence*, 3(2):5–32, 2003.

- [14] Annika Flycht-Eriksson and Arne Jönsson. Dialogue and domain knowledge management in dialogue systems. In *1st SIGdial Workshop on Discourse and Dialogue, Hong Kong*, 2000.
- [15] Annika Flycht-Eriksson, Arne Jönsson, Magnus Merkel, and Håkan Sundblad. Ontology-Driven Information-Providing Dialogue Systems. In *Proceedings of the Americas Conference on Information Systems*, Tampa, Florida, USA, 2003.
- [16] Annika Flycht-Eriksson and Arne Jönsson. Some empirical findings on dialogue management and domain ontologies in dialogue systems – implications from an evaluation of birdquest. In *Proceedings of 4th SIGdial Workshop on Discourse and Dialogue, Sapporo, Japan*, 2003.
- [17] Genevieve Gorrell. Recognition error handling in spoken dialogue systems. In *Proceedings of 2nd International Conference on Mobile and Ubiquitous Multimedia*, 2003.
- [18] Genevieve Gorrell. Using statistical language modelling to identify new vocabulary in a grammar-based speech recognition system. In *Proceedings of Eurospeech 2003, Geneve, Switzerland*, 2003.
- [19] Genevieve Gorrell. Language modelling and error handling in spoken dialogue systems. Licentiate Thesis, Linköping Studies in Science and Technology, Thesis no. 1094,, 2004.
- [20] Genevieve Gorrell, Ian Lewin, and Manny Rayner. Adding intelligent help to mixed initiative spoken dialogue systems. In *Proceedings of ICSLP 2002, Denver, Colorado*, 2002.
- [21] L. Hassling, A. Babic, A. Jönsson, U. Lönn, and H. Casimir-Ahn. Assessing patient information needs as a part of man-machine dialogue development. In S. Bakken, editor, *A Medical Informatics Odyssey: Visions of the Future and Lessons from the Past. Journal of the Medical Informatics Association. Symposium Supplement, AMIA Annual Symposium*, 2001.
- [22] Aseel Ibrahim and Pontus Johansson. Multimodal dialogue systems: a case study for interactive tv. In *Proceedings of the 7th ERCIM Workshop "User Interfaces for All", Paris, France*, 2002.
- [23] Aseel Ibrahim and Pontus Johansson. Multimodal Dialogue Systems for Interactive TV Applications. In *Proceedings of the fourth IEEE ICMI'02 conference, Pittsburgh, PA, USA*, 2002.
- [24] Aseel Ibrahim, Jonas Lundberg, and Jenny Johansson. Speech enhanced remote control for media terminal. In *Proceedings of Eurospeech 2001, Aalborg, Denmark*, 2001.
- [25] Håkan Johansson. Unification-based multimodal integration using integration patterns. In *Proceedings of the 4th Swedish Symposium on Multimodal Communication, Stockholm, Sweden*, 2000.

- [26] Pontus Johansson. Madfilm - a multimodal approach to handle search and organization in a movie recommendation system. In *Proceedings of the 1st Nordic Symposium on Multimodal Communication. Helsingör, Denmark*, pages 53–65, 2003.
- [27] Pontus Johansson. Natural language interaction in personalized epgs. In *UM'03 3rd Workshop on Personalization in Future TV, Pittsburgh, USA*, 2003.
- [28] Pontus Johansson. Design and Development of Recommender Dialogue Systems. Licentiate Thesis 1079, Linköping Studies in Science and Technology, Linköping University, 2004.
- [29] Pontus Johansson, Lars Degerstedt, and Arne Jönsson. Iterative Development of an Information-Providing Dialogue System. In *Proceedings of 7th ERCIM Workshop*, 2002.
- [30] Arne Jönsson and Magnus Merkel. Some issues in dialogue-based question-answering. In *Working Notes from AAAI Spring Symposium, Stanford (To appear)*, 2003.
- [31] Arne Jönsson, Frida Andén, Lars Degerstedt, Annika Flycht-Eriksson, Magnus Merkel, and Sara Norberg. Experiences from Combining Dialogue System Development with Information Access Techniques. In Mark T. Maybury, editor, *New Directions in Question Answering*. AAAI/MIT Press., 2004.
- [32] Håkan Johansson. Understanding multimodal interaction by exploiting unification and integration rule. In *Extended abstract of presentation at the 13th Nordic Conference on Computational Linguistics (NoDaLiDa'01), Uppsala, Sweden*, 2001.
- [33] Staffan Larsson, Lena Santamarta, and Arne Jönsson. Using the Process of Distilling Dialogues to Understand Dialogue Systems. In *Proceedings of 6th International Conference on Spoken Language Processing (ICSLP2000/INTERSPEECH2000), Beijing, China*, 2000.
- [34] Magnus Merkel and Arne Jönsson. Towards multimodal public informations systems. Presented at 13th Nordic Conference on Computational Linguistics, NoDaLiDa '01, Uppsala, Sweden, 2001.
- [35] Magnus Merkel, Michael Petterstedt, and Arne Jönsson. Enhancing access to public information. In *Proceedings of the 7th ERCIM Workshop "User Interfaces for All", Paris, France*, 2002.
- [36] P. Qvarfordt, D. Beymer, and S. Zhai. Realtourist – a study of augmenting human-human and human-computer dialogue with eye-gaze overlay. In *Proceedings of INTERACT, IFIP Conference on Human-Computer Interaction*, 2005.
- [37] Pernilla Qvarfordt. Spoken feedback in multimodal interaction: effects on user experience of qualities of interaction. In P. Paggio, K. Jokinen, and A. Jönsson,

- editors, *Proceedings of the 1st Nordic Symposium on Multimodal Communication, CST Working papers, Report no. 6.*, pages 21–34, 2003.
- [38] Pernilla Qvarfordt. User experience of feedback in multimodal interaction. Licentiate Thesis no. 1003, School of Engineering at Linköping University, 2003.
- [39] Pernilla Qvarfordt. *Eyes on Multimodal Interaction*. PhD thesis, Linköping Studies in Science and Technology, Linköping University, Sweden, Dissertation No. 893, 2004.
- [40] Pernilla Qvarfordt, Arne Jönsson, and Nils Dahlbäck. The role of spoken feedback in experiencing multimodal interfaces as human-like. In *Proceedings of ICMI'03, Vancouver, Canada*, 2003.
- [41] Pernilla Qvarfordt and Lena Santamarta. A first-personnes approach to cooperative multimodal interaction. In T. Tan, Y. Shi, and W. Gao, editors, *Advances in Multimodal Interaction - ICMI 2000*, volume Lecture Notes in Computer Science 1948, pages 650–657. Berlin: Springer, 2000.
- [42] Pernilla Qvarfordt and Shumin Zhai. Conversing with the user based on eye-gaze pattern. In *Proceedings of the Conference on Human-Factors in Computing Systems, CHI 2005, Portland*, pages 221–230. ACM Press, 2005.
- [43] Pontus Wärnestål. Modeling a dialogue strategy for personalized movie recommendations. In *Proceedings of Beyond Personalization 2005 Workshop (Intelligent User Interfaces'05). San Diego (CA)*, pages 77–82, 2005.
- [44] Pontus Wärnestål. Modularized user modeling in conversational recommender systems. In L. Ardissono, P. Brna, and A. Mitrovic, editors, *Proceedings of 10th International Conference on User Modeling (UM05) LNAI 3538. Edinburgh, Scotland U.K.*, pages 545–547. Springer-Verlag, Berlin Heidelberg, 2005.
- [45] Pontus Wärnestål. User evaluation of a conversational recommender system. In *Proceedings of the 4th Workshop on Knowledge and Reasoning in Practical Dialogue Systems (International Joint Conference on Artificial Intelligence 2005). Edinburgh, Scotland U.K.*, pages 32–39, 2005.