IEEE CIS

CDS NEWSLETTER

The Newsletter of the Technical Committee on Cognitive and Developmental Systems

Developmental Robotics
Machine Intelligence
Neuroscience
Psychology

Volume 16, number 1
2019

Editorial

Leveraging Developmental Sciences for Impactful Educational Technologies

Pierre-Yves Oudeyer

Inria and Ensta
ParisTech, France
Editor

pierre-yves.oudeyer@inria.fr

Fundamental research aiming to understand better how children learn and develop can have a major societal impact. One key example is education: in the last decades, several advances enabled to show how certain learning and teaching techniques could improve significantly comprehension and memorization in children. A major challenge consists in translating these advances into real classroom practices, a challenge at the core of "translational educational sciences". In this newsletter, a dialog initiated by Georges Kachergis highlights how adaptive learning technologies, e.g. educational apps, can be both an efficient channel for this translation, and an opportunity for further understanding how to foster efficient learning in the classroom. Several experts of this domain provide their point of view: Jennifer Zosh, Roberta Golinkoff, Kathy Hirsch-Pasek, Rebecca Dore, Brenna Hassinger-Das, Benedict du Boulay and Ken Koedinger. They discuss various challenges to be addressed, ranging from establishing strong collaborations between developmentalists, app developers and teachers, to deploying large-scale ecologically valid experimentations. They also offer a perspective on how artificial intelligence can play a key role in impactful educational apps, enabling to implement personalized and motivating learning strategies.

Then, a new dialog initiation is proposed by Clément Moulin-Frier on a fundamental question for the future of AI: How to evaluate open-ended learning agents? While impressive progress was made recently in reinforcement learning, as shown by performance in benchmarks with a set of pre-defined external objectives, a major challenge is now to build autonomous agents that can progressively discover and learn open repertoires of skills in open worlds. Several steps in this direction have been made in the last decade, such as algorithms enabling agents to imagine their own goals and self-organize their learning curriculum (initially in the CDS community, and recently extended in the machine learning community). Another related line of work is open-ended evolution in multi-agent systems with co-evolution. However, a key challenge is how to measure progress in this area, as traditional RL benchmarks were not constructed to address open-ended learning. I invite all readers interested to participate to this dialog to send me their response by June 30th, 2020. The length of each response must be between 600 and 800 words including references (contact pierre-yves.oudeyer@inria.fr).

Table of Contents

Editorial

Dialogue

New Dialogue Initiation

IEEE TCDS Table of Contents

Dialogue

Leveraging Adaptive Games to Learn How to Help Children Learn Effectively

George Kachergis

Dept. of Artificial Intelligence
Donders Institute for Brain, Cognition, and Behaviour
Radboud University

g.kachergis@donders.ru.nl

Despite more than a century of rigorous psychological research on learning and memory which has discovered robust methods for improving learning, relatively few large shifts in education have come about due to insights gained about how people learn (Roediger, 2013). With an eye towards changing this, Dunlosky et al. (2013) reviewed ten of the most well-studied learning techniques that are ripe for translation and found that the most effective and generally useful techniques for a variety of materials and learning contexts are: using spaced (i.e. distributed) practice as opposed to massed repetitions, retrieval practice (aka testing), and interleaved practice—mixing topics of study over short periods. While these principles might be difficult to apply in a standard classroom, with only a single teacher to facilitate interleaved study and retrieval practice, software has the potential to adapt to individual learners, recording responses and adjusting the content based on learning progress. Indeed, research on computer-based cognitive tutors has promoted the translation of memory and learning research into educational practice in at least some classrooms, and such systems have proven effective (for a review, see Koedinger & Corbett, 2005).

However, with the growing ubiquity of tablets and smartphones, software developers are already offering tens of thousands of educational games and applications on the app stores. While parents often flock to the well-designed apps that may promise to teach basic literacy or numeracy skills and that are most engaging and entertaining for their children, the bulk of these apps do not draw on principles from the science of learning (see Hirsh-Pasek et al. 2015 for a review). The main question of this dialogue initiation is thus, "How can we work with app developers--or develop our own apps--to help translate our research into educational practice, while also using data collected from these apps to learn more about development?" The challenges can be organized into a few categories: 1) app development, 2) organizational, and 3) scientific.

Software Challenges

Most researchers do not have the resources to hire an app developer, and the reverse is also generally the case. As researchers, can we reach out to app developers to advise on learning principles they may want to incorporate? Would app developers then agree to help us in collecting data to enable research and further improvements? A few edTech startups have been built with this reciprocal arrangement in mind, with both researchers and app developers on board (e.g., Cognitive Toybox and egoTeach, my own attempt), and can find support from industry and other funding sources that are not possible without an industry partner. I also encourage researchers to look at popular education apps, evaluate their scientific basis (for examples, see Hirsh-Pasek et al. 2015) and to reach out to app developers and discuss collaboration.

Organizational Challenges

When a researcher is faced with deploying an app for testing, there are diverse approaches that result in distinct types of datasets. Deploying in the competitive environment of the app stores will often result in slow (or no) adoption and perhaps not very extended use, unless the app is quite engaging. Even a large naturalistic dataset can present its own difficulties: most users contribute little data, so usage- and age-matched comparisons can be a challenge and long learning trajectories are hard to come by. Aside from recruiting and running individual participants for short-term studies, are there ways to engage schools for larger-scale, longer-term testing? What types of apps and content would teachers most welcome? How can we work together as teams of researchers along with educators to create more universal learning aids that also allow us to gather data?

Scientific Challenges

Aside from further testing strategies such as distributed practice and retrieval practice or investigating new techniques, there is still much to be learned about which methods can be synergistically combined, and which might interfere with each other. For example, some recent work has investigated the relationship between curiosity, motivation, and learning in educational apps (Oudeyer, Gottlieb, & Lopes, 2016). Which learning techniques can be implemented and studied jointly in an app? Another issue is how to measure generalization of the skills outside the app--in the classroom, or in other apps. Once there are multiple researchers field testing, say, numeracy apps, how can they coordinate recruitment to ensure that they have different participants? With a population of diverse participants on the app stores, with varied backgrounds and interests, it is also interesting to consider if individual differences in cognitive abilities might influence the effectiveness of some strategies, as well as participants' behavior and motivation.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. Psychological Science in the Public Interest, 14(1), pp. 4–58.

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., and Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. Psychological Science in the Public Interest, 16(1), pp. 3–34.

Koedinger, K., & Corbett, A. (2005). Cognitive Tutors. In R. Sawyer (Ed.), The Cambridge Handbook of the Learning Sciences (Cambridge Handbooks in Psychology, pp. 61-78). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511816833.006

Oudeyer, P.-Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. Progress in Brain Research, 229, pp. 257–284.

Roediger, H. L. (2013). Applying cognitive psychology to education: Translational educational science. Psychological Science in the Public Interest, 14(1), pp. 1–3.

Leveraging the How and What of Learning and App Development to Create Truly Educational Apps and Move Science Forward

Jennifer M. Zosh

The Brandywine Child Development Lab, Penn State Brandywine University, USA

jmz15@psu.edu

Roberta M. Golinkoff

Dept. of Psychology and Linguistics, University of Delaware, USA

roberta@udel.edu

Kathy Hirsh-Pasek

Department of Psychology, Temple University, Brookings Institute, USA

khirshpa@temple.edu

Kachergis' Dialogue Initiation provides a comprehensive view of the landscape of today's educational app terrain. The crux of the challenge is that app developers and developmental scientists live in two parallel universes that rarely touch base with one another (Dore et al., 2018; Hirsh-Pasek, Zosh et al., 2015). Business as usual will not help us develop top educational apps but forming bi-directional partnerships and a shared mission will. We suggest that one way forward is to have a shared framework centered around evidence-based principles based in the science of learning. Framing dialogue between developers and scientists within the context of how to create educational apps and what children need to learn for future success will yield the very best educational apps for children.

The How of learning?

Hirsh-Pasek, Zosh, and colleagues (2015) suggest that the science of learning literature supports that children learn best when they are: 1) active with the content ("minds on," not just tapping and swiping); 2) engaged (not distracted by peripheral "bells and whistles"); 3) learning meaningful information (e.g., learning that a triangle isn't just the shape on the screen but also approximates the shape of a pizza slice) and critically; 4) benefiting from socially interactive experiences around the app with peers and adults. Apps can be designed in ways to leverage the pedagogical benefit of these pillars—or, not. Apps can promote active learning when children are asked to think deeply and engage with a learning goal and not just tap until they happen upon the correct answer. Designers should be mindful of the features added to apps and ask whether the pop-up game or fun graphic supports or distracts from minds on thinking. Apps can take learning beyond the screen and engage children in thinking how their learning manifests in the outside world and not just within the app itself. Finally, new technology can be leveraged in ways that promote social interaction rather than isolation. Apps have the potential to help learning come alive and be truly deserving of the "educational" classification, but only when they support these principles rather than act as digital worksheets.

The What of Learning

In our 2015 piece, we imagined a new frontier of development in which the pillars aligned more closely with app design. Yet, educational apps tend to approach education as if delivery of content is key. This results in apps reinforcing correct answers as a staple of the "educational" experience. There is no doubt that content is important. An app with a learning goal focused on fractions needs to understand why children find ½ so much easier to interpret than 3/8. But learning involves so much more than memorizing content. Books like Golinkoff and Hirsh-Pasek's (2016) Becoming Brilliant or Dintersmith's (2018) What School Could Be illustrate how children's learning relies on a suite of skills that allow for, but go beyond mere memorization of content. Golinkoff and Hirsh-Pasek (2016) call these skills that 6 C's: Collaboration, Communication, Content, Critical Thinking, Creative Innovation, and Confidence. In combination, the use of these 6C's allow for app development that fosters discovery learning rather than rote learning and enrich learning to learn skills while promoting content at the same time. Indeed, a number of research projects now speak to the advantage of guided play approaches to learning that encourage active discovery-based learning (Zosh et al., 2018). As Dintersmith (2018) writes, "remarkable things happen when kids learn organically and passionately, instead of painting by number" (p. 43). Apps need to move beyond being digital flash cards to offer deep and meaning experiences that embrace a range of 21st Century Skills.

The What and How of Educational Apps: Promoting learning and new scientific knowledge

Kachergis speaks to a new potential for research in which he talks of large-scale data collection across groups of learners with various background knowledge and abilities. However, we must ensure that the apps being studied are of the highest quality. And high quality learning apps are contingent on our drawing from two cultures of developmentalists and developers who can ensure that the products under the microscope are built with the best elements of the how and what of learning and app development. By answering questions about how children tend to use apps, we may begin to uncover new insights about the mind. For example, collecting large data sets that uncover the common mistakes children make in the study of fractions might elucidate the features that are problematic in learning fractions. By studying what children do while using the app, we can answer new questions about how to engage children in learning. What are the elements that keep children engaged and not distracted? What technologies are available to help children break down the screen barrier, engage in meaningful learning, and take the lessons from the app to real world problems? What are children clicking on, when, and how does their performance change with the introduction of a new feature?

The potential is clear – if only two cultures with a shared mission can work together to create apps that confer genuine learning opportunities for children.

Dintersmith, T. (2018). What school could be: Insights and inspiration from teachers across America. Princeton, NJ: Princeton University Press.

Dore, R. A., Shirilla, M., Verdine, B. N., Zimmermann, L., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Developer meets developmentalist: improving industry–research partnerships in children's educational technology. Journal of Children and Media, 12(2), 227–235. http://doi.org/10.1080/17482798.2018.1450086

Golinkoff, R. & Hirsh-Pasek, K. (2016). Becoming brilliant: What science tells us about raising successful children. Washington, DC: APA LifeTools.

Hirsh-Pasek, K.*, Zosh, J.M.* (*joint first authors), Golinkoff, R. M., Gray, J., Robb, M., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the Science of Learning. Psychological Science in the Public Interest, 16, 3-34.

Zosh, J.M., Hirsh-Pasek, K., Hopkins, E, Jensen, H. Liu, C., Neale, D., Solis, S.L., & Whitebread, D. (2018). Accessing the inaccessible: Play as a spectrum. Frontiers in Psychology, 9.

Translating Research Principles into Effective Partnerships

Rebecca A. Dore

School of Education, Ohio State University, USA
rdore@udel.edu

Brenna Hassinger-Das

Psychology Department, Pace University, USA
bhassingerdas@pace.edu

Kathy Hirsh-Pasek

Department of Psychology, Temple University, Brookings Institute, USA
khirshpa@temple.edu

Roberta M. Golinkoff

Dept. of Psychology and Linguistics, University of Delaware, USA
roberta@udel.edu

The question of how to effectively translate learning science principles into education is even broader than developing evidence-based educational technologies. For too long, schools have followed outdated learning theories that offer "drill and kill" and "fill in the bubble" testing. The advent of new educational technologies must do better than what we currently offer else we merely translate current educational practices to technological platforms. We thank Kacergis for raising a number of important issues.

First, we respond to Kachergis's category of "app development". One challenge is that researchers and developers talk past each other and fail to see the barriers that prevent clear communication. As we write in our recent piece about connecting developmentalists and developers in the Journal of Children and Media:

> The technical languages of our respective fields are laced with jargon, so do not assume everyone will understand and don't hesitate to ask clarifying questions. For example, in one of our partnerships, the word "assets" was used continually by the developers to refer to art and audio for the app... Eventually both groups caught on...but avoiding jargon and clearly defining terms can help reduce the mutual learning curve (Dore at al., 2018, p. 230).

We urge readers to consult this piece for more thoughts on establishing research-industry partnerships and avoid some of the pitfalls we initially encountered.

Through several research projects, we have had first-hand experience creating educational technologies that are both empirically-based and easy to use for practitioners and parents. For example, in one project, we worked with an educational app development company and asked them to use their existing app structure to create a new game for our study (Dore, Shirilla et al., in preparation). In exchange, we would provide data on the effectiveness of the game that would be useful to their marketing. However, given the nature of data collection and analysis, it was over a year from the time we received the game to when we provided a report of its educational effectiveness. By that time, the company was in the process of changing their business model and the information was no longer as helpful as it would have been previously.

Second, within the category of organizational challenges, Kachergis offers a most innovative and exciting idea: to develop collaborations with schools so that educational apps can be tested iteratively. With the right research-practice partnership, one that truly favors the joint creation of educationally appropriate materials, a school could serve as a valuable laboratory for testing hypotheses and creating apps that speak to specific school needs. Undoubtedly apps created in this way would have applicability beyond the specific setting in which they were developed. However, we also recognize that this path is often fraught with obstacles, and requires that schools, app developers, and scientists join forces--and not in superficial ways. But imagine the products that could emerge from such partnerships! And the novel approaches that could result from combining the use of apps with classroom activities as well as spurring learning outside classroom walls.

In Kachergis's third category, we acknowledge the scientific challenges that face researchers who wish to create worthwhile apps according to principles from the science of learning (Hirsh-Pasek, Zosh, Golinkoff et al., 2015). These principles, grounded in thousands of research papers, are instructive for app development and evaluation the first principle is that we learn best when we are active rather than passive. Apps should promote a "minds-on" approach in which the learner has to mentally manipulate ideas. Second, we learn best when we are engaged rather than distracted. Effective educational technology ensures that learners are involved with the learning material itself and not distracted by peripheral elements that might appear. Third, we learn best in meaningful contexts that connect to children's lives outside of the screen. Finally, we learn best through social interaction. Apps and educational technology that encourages connections between students and their teachers and peers is often more effective than technology that fosters app use in individual silos. Regardless, even the best apps will not be a panacea but should be seen as part of an overall educational approach focused on guided discovery and hands-on learning (Yu et al., 2018).

Lastly, dissemination can take many forms. Researchers from our group have given talks at conferences where developers congregate, like the Dust or Magic Institute. In addition to publishing in academic journals, researchers would do well to write pieces in publications intended for practitioners and lay audiences such as the American Educator. Finally, websites like Common Sense Media already have thoughtful and research-informed reviews of popular apps and digital games. If researchers can partner with similar organizations, translating and disseminating relevant evidence to lay audiences may be within reach.

Dore, R. A., Shirilla, M., Hopkins, E. J., Collins, M., Scott, M., Schatz, J., S, Lawson-Adams, J., Valladares, T., Foster, L., Puttre, H., Toub, T. S., Hadley, E., Golinkoff, R., Dickinson, D. & Hirsh-Pasek, K. (in preparation) Education in the app store: Using a mobile game to support preschoolers' vocabulary learning.

Dore, R. A., Shirilla, M., Verdine, B. N., Zimmermann, L., Golinkoff, R. M., Hirsh-Pasek, K. (2018). Developer meets developmentalist : improving industry – research partnerships in children's educational technology. Journal of Children and Media, 2798, 1–9. https://doi.org/10.1080/17482798.2018.1450086

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. Psychological Science in the Public Interest, 16, 3–34, doi:10.1177/1529100615569721.

Yu, Y., Shafto, P., Bonawitz, E., Yang, S. C. H., Golinkoff, R. M., Corriveau, K. H., ... & Xu, F. (2018). The theoretical and methodological opportunities afforded by guided play with young children. Frontiers in Psychology, 9.

We Need to Distinguish Conceptual Learning from Reinforcement Learning

Benedict du Boulay

Human Centred Technology, University of Sussex, United Kingdom
b.du-boulay@sussex.ac.uk

I agree with the general thrust of Kachergis' argument that many educational applications have failed to take account of what is known about human learning and that the barriers to improving this situation are multiple. Let me start with my main point. This problem is recognised and something is being done, at least in the UK, see the Educate Project (e.g., https://www.ucl.ac.uk/ioe/departments-and-centres/centres/ucl-knowledge-lab/six-research-themes/designing-smart-technologies-teaching-and-l/educate). The Educate Project, based at the Knowledge Lab, University College London, provides training to educational application developers to help them understand more about the science of learning as well as about how to conduct educational evaluations. Evaluation can be applied at all stages of the design and development cycle. Such evaluations range from paper or soft prototyping using learner-centred techniques to detailed comparative evaluations of the application in authentic educational contexts.

The paper by Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013) provides a useful, well-researched checklist of techniques that can be applied by students (and teachers) to achieve more effective learning. However, there are two aspects to the term "learning": one focusing on comprehension and understanding, and the other on consolidation of that understanding and memorisation. The techniques identified by Dunlosky et al. that worked best were essentially about memorisation and consolidation (e.g. the use of spaced practice) as opposed to initial comprehension and understanding. By contrast, Dunlosky at al. did not find sufficient evidence to support the use of two comprehension strategies, namely self-explanation and elaborative interrogation, despite their intuitive appeal. A recent meta-analysis of self-explanation is more positive and found a mean effect size of 0.55 across 69 comparative studies of self-explanation vs. other techniques (Bisra, Liu, Nesbit, Salimi, & Winne, 2018). These authors also make the following suggestion: "Due to the limitations of relying on instructor-scripted prompts, we recommend that future research explore computer generation of self-explanation prompts."

The above suggestion by Bisra et al. takes us into the realm of Artificial Intelligence in Education. This field specializes (and has specialized since the 1970s) in developing systems to assist comprehension and understanding. It includes the Cognitive Tutors (Koedinger & Corbett, 2005) mentioned by Kachergis, and also covers other kinds of learning environment based on research into learning. Such environments support for example, learning by teaching (Biswas, Segedy, & Bunchongchit, 2016), peer tutoring (Walker, Rummel, & Koedinger, 2009), metacognition (Trevors, Duffy, & Azevedo, 2014), self-explanation (Conati & Van Lehn, 2000) and elaborative interrogation (Nye, Graesser, & Hu, 2014; VanLehn et al., 2005).

In addition, over the last few years there have been a number of meta-analyses comparing intelligent tutoring systems and intelligent learning environments against conventional classrooms and one-to-one human tutoring. The overall message is that, compared to conventional classrooms, intelligent tutoring systems and intelligent learning environments do better with an effect size of 0.47, but do worse than one-to-one human tutoring with an effect size of 0.19 (du Boulay, 2016).

One of the issues concerning educational application development, not mentioned by Kachergis, is that building applications to support memorisation and consolidation is less complex and cheaper than those to support comprehension and understanding, so inevitably the market favours the former. However, the big gains will be via developing the latter and that is where our efforts should be focused. It is also important to note that the biggest barrier to learning in rich societies is not the technology, but the motivation and values of the learners, who often find formal learning boring, frustrating and without obvious value to them.

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing Self- Explanation: a Meta-Analysis. Educational Psychology Review, 30(3), 703-725.
Biswas, G., Segedy, J. R., & Bunchongchit, K. (2016). From Design to Implementation to Practice a Learning by Teaching System: Betty's Brain. International Journal of Artificial Intelligence in Education, 26(1), 350-364.
Conati, C., & Van Lehn, K. (2000). Toward Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation International Journal of Artificial Intelligence in Education, 11(4), 389-415.
du Boulay, B. (2016). Artificial Intelligence As An Effective Classroom Assistant. IEEE Intelligent Systems, 31(6), 76-81.
Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. Psychological Science in the Public Interest, 14(1), 4-58.
Koedinger, K. R., & Corbett, A. (2005). Cognitive Tutors. In R. K. Sawyer (Ed.), The Cambridge Hanbook of the Learning Sciences (pp. 61-78): Cambridge University Press.
Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and Family: A Review of 17 Years of Natural Language Tutoring. International Journal of Artificial Intelligence in Education, 24(4), 427-469.
Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. Educational technology research and development, 62(5), 507-528.
VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., . . . Wintersgill, M. (2005). The Andes Physics Tutoring System: Lessons Learned. International Journal of Artificial Intelligence in Education, 15(3), 147-204.
Walker, E., Rummel, N., & Koedinger, K. R. (2009). Integrating Collaboration and Intelligent Tutoring Data in the Evaluation of a Reciprocal Peer Tutoring Environment. Research and Practice in Technology Enhanced Learning, 4(3), 221- 251.

Ken Koedinger

Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, USA
koedinger@cmu.edu

Learning Technology Facilitates Research-Practice Partnerships Toward Better Education

We connect to the dialog initiative by summarizing elements of our prior work toward achieving "translational educational sciences" particularly though "cognitive science principles used in real-world large-scale educational technologies" and often employing "ecologically valid educational technology experiments".

Many organizations now operate online or blended learning environments with large numbers of learners (e.g., Khan, EdX, Carnegie Learning, Duolingo, Coursera, ASSISTments, Florida Algebra Nation, Enlearn, Open Learning Initiative, etc.). These environments have the potential to (1) transform the way that learning science is conducted, (2) create an infrastructure for continuous improvement, and (3) accelerate learning so students achieve better outcomes at rates well beyond current practice. Yet, a huge "opportunity chasm" resides between the current poor state of education, on one hand, and the tremendous research-demonstrated potential of learning science and technology, on the other. So, what's missing? Why have we not seen more progress in addressing achievement gaps or reducing costs? Why is so little of this great science and technological potential not making a difference in schools?

We assert the translational problem is not fundamentally about implementation or application of learning science principles. Instead, the fundamental challenge is to develop a practical learning science that optimizes instructional methodologies in real educational settings. We need a shared intellectual capacity and institutional commitment among interdisciplinary teams of scientists, engineers, and instructors to make use of technology-enabled, data-driven methods for continuous improvement. The Knowledge-Learning-Instruction (KLI) framework (Koedinger, Corbett, & Perfetti, 2012) was designed to assist with moving learning research and theory to practice by recognizing the import of data in instructional design decisions and advocating for using data to advance learning theory.

A growing number of cases indicate that such interdisciplinary teams can design and develop technology-enabled, blended educational solutions. It is worth emphasizing that "blended" means these solutions not only incorporate advanced technology, but also address the surrounding social context, especially the role of the teacher and of peer learners. Large-scale, highly rigorous randomized field trial experiments have demonstrated the potential of these science and technology-enhanced learning solutions to dramatically accelerate learning: double high school algebra achievement in a school year (Pane et al., 2016), produce much better college statistics learning in half a semester (Lovett, Meyer, & Thille, 2008), significantly enhance middle school math learning (Roschelle et al., 2016).

To be so highly effective, these solutions are adapted to particulars of the domain content and thus straightforward generalization to new content does not reproduce the same accelerated learning. While the solutions cannot be copied, the process can. All these cases have involved interdisciplinary teams engaging in a multi-year iterative engineering process of continually improving the technology and surrounding social processes. The full potential of science and technology-enhanced learning may be well beyond the demonstrations of 2x acceleration of learning as stated above. Two examples make the case for how learning can be accelerated through a combination of iterative design based on the psychology of learning, informed analytics of student data, and technology-enhanced implementation. One, early experiments with Cognitive Tutors demonstrated students learn programming better with a Cognitive Tutor in one-third the time it takes to learn programming without one and demonstrated about one letter grade improvement in learning high school geometry (Anderson, Corbett, Koedinger & Pelletier, 1995). Two, contrary to teacher beliefs, data-driven discoveries found that beginning algebra students have less difficulty solving story problems than matched equations. More than a dozen classroom-based experiments have tested the effect of blended use of the Cognitive Tutor Algebra course, the largest of which involved random assignment of over 18,000 students in 147 schools across 7 states. Over the school year, students in the Cognitive Tutor classes learned twice as much, as measured by an independent standardized math test, as those in the regular algebra classes (Pane et al., 2016).

A reliable doubling of learning rate would be revolutionary and have clear economic implications. The demands of interdisciplinary data-driven design and engineering are substantial, but new evidence suggests that well engineered courses may accelerate learning well beyond the two examples above. We have analyzed data from 10,000+ students using OLI courses as part of MOOCs or credit-bearing classes. We find that students differ widely in their use of learning resources and these differences are associated with huge differences in learning outcomes and the time it takes to achieve them. Students who spend more time doing online questions and problems with automated feedback and instruction learn about five to ten times more than students who spend the same amount of time reading online text or watching online lecture videos (Koedinger et al., 2015). Experiments are underway to see if encouraging more doing, as well as more broadly addressing other opportunities to enhance student meta-cognition, motivation, and participation (Koedinger, Aleven, Roll, & Baker, 2009), can achieve learning acceleration that approaches this 10x difference. Internet-scale experiments involving 10s or 100s of thousands of students provide a path to such order of magnitude improvement in education.

We agree with dialog initiative in suggesting much more can be done for researchers to engage educational technology vendors, and this need not be

limited to educational games. It is worth highlighting that the current state of affairs is that such technologies are not just in "some classrooms" but are in many classes. As such, there is huge opportunity for researchers to collaborate with the many developers, commercial and university-based, who have online or blended learning environments in use by vast numbers of learners.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. The Journal of the Learning Sciences, 4(2), 167-207.

George Kachergis

Dept. of Artificial Intelligence
Donders Institute for Brain, Cognition, and Behaviour
Radboud University

g.kachergis@donders.ru.nl

Summary of Responses to Dialog Initiation "Leveraging Adaptive Games to Learn How to Help Children Learn Effectively"

In this issue of the IEEE CIS Newsletter on Cognitive and Developmental Systems, seven researchers from psychology, human-computer interaction, education, and cognitive science respond to my position statement about how to build and sustain bridges between developers and scientists to start a virtuous cycle of improving educational practice as well as our understanding of how people learn. All of the respondents agreed that there is ample room for more engagement between researchers and educational technology developers, highlighting and elaborating upon various barriers to these collaborations. Encouragingly, some pointed out successful past and ongoing collaborations (Ken Koedinger, for example) as well as tips on establishing successful research-industry partnerships (Rebecca A. Dore and colleagues).

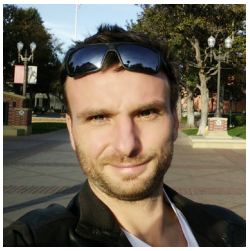
The responses had a variety of recommendations for how to sustainably bridge the divide between researchers and developers. Benedict du Boulay describes an ongoing research outreach effort at University College London that teaches developers of educational apps about the science of learning, and how to perform educational evaluation. Jennifer M. Zosh and colleagues propose that bi-directional partnerships can best be established around a shared educational goal if researchers clearly frame the dialogue with app developers around how principles from the science of learning can be implemented in apps, as well as separately identifying the content and skills that should be included. Rebecca A. Dore and colleagues have found that working directly with app developers can be difficult due to the different timescales at which research and industry operate, but are optimistic that serious, long-term collaborations between researchers, educators in schools, and (optionally, perhaps) app developers can iteratively improve both the apps and our science of learning. Ken Koedinger echoes this, even arguing that it is more representative of the fundamental problem: developing and deploying a practical learning science in real educational settings and optimizing it over years requires commitment from institutions and interdisciplinary teams including not only scientists and engineers, but also instructors. I agree that this is one proven model, but its scarcity leads me to suggest it is difficult to emulate and to deploy at scale, given the necessary commitment.

All responses emphasized to some degree the importance—and capability—of using educational technology to broaden the social context of learning, engaging learners with their peers and teachers. Rebecca A. Dore and colleagues emphasize that technology should enable us to go beyond the out-of-date standard education methods still widely-used to deliver content, and give learners a chance to learn actively, in meaningful, engaging (and not distracting) contexts. Jennifer M. Zosh and colleagues highlight that technology should be used to go beyond teaching classic content, and extend to 21st century skills such as collaboration and critical thinking. This is echoed and further nuanced by the distinction Benedict du Boulay makes between two facets of learning: comprehension and understanding vs. consolidation and memorization. Benedict du Boulay points out that it is more complex and thus costly to build software that supporting initial comprehension and understanding (e.g., peer tutoring and learning by teaching) than simpler applications supporting memorization, which are thus favored by the market. Ken Koedinger stresses that technology is only one part of the blended educational solutions that would be most effective, which must include roles for teachers and peer learners—and which promise to offer remarkable learning gains.

A final theme that came up in some of the responses was the fact that learners in different contexts—social, economic, and beyond—vary tremendously in their motivation, values, and chosen strategies for pursuing formal learning. Benedict du Boulay notes the diverse barriers to learning faced by different populations: learners in rich societies may not lack the technology, but rather the motivation and values to pursue formal learning. Ken Koedinger identifies that some learners consistently choose more effective study strategies (e.g. practice problems) than others (reading and watching lectures), and hypothesizes that intervention studies may help learners learn more effectively. It is certainly true that learners have diverse motivations and strategies, and it is reasonable to expect that we will need further data from a variety of learners in diverse educational contexts in order to improve both education and the science of learning. We have identified some obstacles in building research-industry partnerships, but have also identified hopeful examples and a unified goal for all partners.

New Dialogue Initiation

How to Evaluate Open-ended Learning Agents?



Clément Moulin-Frier

Research Scientist
at Cogitai,
California, USA

clement.moulinfrier@gmail.com

Understanding and modeling the mechanisms supporting open-ended learning in biological and artificial agents is a major long-term objective of the Cognitive and Developmental Systems community (CDS). On the modeling side, the field of Machine Learning (ML) has recently made important steps in this direction. These advances call for new methods evaluating the ability of learning agents to autonomously discover and learn new tasks of increasing complexity[1].

The integration of deep neural networks with reinforcement learning algorithms (deep reinforcement learning, DRL, Mnih et al., 2015) has provided a powerful framework for implementing agents learning specific tasks in complex environments. Those tasks are defined by a reward function that the agent aims at maximizing over the long-term, by learning an action policy mapping its observations (possibly in high-dimensional spaces such as raw images) to its actions. Interestingly, some principles originally formalized by the CDS community have played an important role in some of the recent extensions of this framework. Intrinsically motivated learning, goal babbling and curriculum learning are now widely used in DRL experiments to solve complex problems (e.g. Andrychowicz et al., 2017; Pathak, Agrawal, Efros, & Darrell, 2017) (in particular sparse-reward problems, where an initial random action policy does not provide informative feedback about the task to be learned). DRL algorithms have also been extended to handle the learning of multiple tasks in parallel, to transfer previously acquired knowledge to new tasks, or to integrate planning mechanisms (e.g. Srinivas, Jabri, Abbeel, Levine, & Finn, 2018).

Altogether, these advances bring us a few steps closer to the possibility of implementing artificial agents capable of open-ended learning, i.e. able to autonomously discover and learn new tasks of increasing complexity. Thus, advances in DRL algorithms have encouraged the ML community to invent more complex evaluation domains, often based on sandbox 3D simulation environments embedding realistic physics (e.g. Beattie et al., 2016), in which challenging tasks are defined (e.g. tasks with sparse rewards or requiring planning abilities). Moreover, some of those environments allow multi-agent simulation and are designed to support the persistence of the world's elements for supporting long-term agent development and lifelong learning (Johnson, Hofmann, Hutton, & Bignell, 2016; Suarez, Du, Isola, & Mordatch, 2019). How can we properly evaluate open-ended learning algorithms in such environments? Is the standard concept of a benchmark even relevant at all for evaluating open-ended learning agents?

Are task-oriented evaluation methods still applicable to open-ended learning agents?

In reinforcement learning, standard evaluation methods are task oriented. We define a domain (an environment with its transition dynamics) and we evaluate the ability of an agent to learn how to solve some predefined tasks in this domain, by measuring e.g. its performance or its convergence time. This method is however hardly applicable to evaluate open-ended learning agents since their objective is precisely to discover their own tasks in an autonomous manner. For this reason, alternative methods have been proposed such as the qualitative analysis of developmental trajectories (Oudeyer, Kaplan, Hafner, & Whyte, 2005), generalized environments for empirical evaluations (Whiteson, Tanner, Taylor, & Stone, 2011) or measures of adaptive progress in co-evolutionary simulations (Cliff & Miller, 1995). Can we still propose task-oriented methods for evaluating open-learning agents? (e.g. by approximating an integral over the space of all possible tasks in some challenging domain). How to define standard metrics of complexity and similarity in continuous task spaces? Can we propose alternative evaluation methods that are not task-oriented?

How to evaluate open-ended co-adaptation in multi-agent environments?

The recent advances in DRL have revived the field of multi-agent reinforcement learning, where several agents are interacting in a shared environment and individually learning how to maximize their own cumulative reward (Littman, 1994). When those rewards are coupled, cooperative or competitive strategies can emerge at the population level, possibly entering a positive feedback loop of coadaptation fostering the co-acquisition of increasingly complex behaviors (Al-Shedivat et al., 2018; Pérolat et al., 2017). The evaluation of such complex systems is challenging because the performance of an individual agent depends on the behaviors of the others, which are themselves adapting through time. How to measure open-ended co-adaptation in multi-agent environments? Can we adapt methods used in Artificial Life for characterizing the dynamics of complex systems continuously evolving through time? What is the role of the massively multi-agent nature of the biological world in the evolution and development of open-ended learning abilities?

1 In this dialog, we will use "open-ended learning" as a general term referring to the capacity of an agent to autonomously discover and learn new tasks of increasing complexity. Related terms have been used in the literature, e.g. continual learning or lifelong learning. see e.g. (Banzhaf et al., 2016; Doncieux et al., 2018; Oudeyer, Kaplan, & Hafner, 2007; Ring, 2005; Silver, Yang, & Li, 2013).

Al-Shedivat, M., Bansal, T., Burda, Y., Sutskever, I., Mordatch, I., & Abbeel, P. (2018). Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments. In International Conference on Learning Representations.

Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, **R., Welinder, P., ... Zaremba, W.** (2017). Hindsight experience replay. In Advances in Neural Information Processing Systems (pp. 5048–5058).

Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R.,

Foster, J. A., McMullin, B., ... White, R. (2016). Defining and simulating open-ended novelty: requirements, guidelines, and challenges. Theory in Biosciences, 135(3), 131–161. https://doi.org/10.1007/s12064-016- 0229-7
Beattie, C., Leibo, J. Z., Teplyashin, D., Ward, T., Wainwright, M., Küttler, H., ... Others. (2016). DeepMind Lab. ArXiv Preprint ArXiv:1612.03801.
Cliff, D., & Miller, G. F. (1995). Tracking the red queen: Measurements of adaptive progress in co-evolutionary simulations. In European Conference on Artificial Life (pp. 200–218).
Doncieux, S., Filliat, D., Díaz-Rodríguez, N., Hospedales, T., Duro, R., Coninx, A., ... Sigaud, O. (2018). Open-Ended Learning: A Conceptual Framework Based on Representational Redescription. Frontiers in Neurorobotics, 12, 59. https://doi.org/10.3389/fnbot.2018.00059
Johnson, M., Hofmann, K., Hutton, T., & Bignell, D. (2016). The Malmo Platform for Artificial Intelligence Experimentation. International Joint Conference on Artificial Intelligence (IJCAI), 4246–4247.
Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. Machine Learning Proceedings 1994, 157–163. https://doi.org/10.1016/B978-1-55860-335- 6.50027-1
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529– 533. https://doi.org/10.1038/nature14236
Oudeyer, P.-Y., Kaplan, F., & Hafner, V. (2007). Intrinsic Motivation Systems for Autonomous Mental Development. IEEE Transactions on Evolutionary Computation, 11(2), 265–286.
Oudeyer, P.-Y., Kaplan, F., Hafner, V. V, & Whyte, A. (2005). The playground experiment: Task- independent development of a curious robot. In Proceedings of the AAAI Spring Symposium on Developmental Robotics (pp. 42–47).
Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self- supervised prediction. In International Conference on Machine Learning (ICML) (Vol. 2017).
Pérolat, J., Leibo, J. Z., Zambaldi, V., Beattie, C., Tuyls, K., & Graepel, T. (2017). A multi-agent reinforcement learning model of common-pool resource appropriation. In Advances in Neural Information Processing Systems 30 (NIPS 2017) (pp. 3643–3652).
Ring, M. B. (2005). Toward a formal framework for continual learning. In NIPS Workshop on Inductive Transfer, Whistler, Canada.
Silver, D. L., Yang, Q., & Li, L. (2013). Lifelong machine learning systems: Beyond learning algorithms. In 2013 AAAI spring symposium series.
Srinivas, A., Jabri, A., Abbeel, P., Levine, S., & Finn, C. (2018). Universal Planning Networks. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden (pp. 4732--4741). PMLR.
Suarez, J., Du, Y., Isola, P., & Mordatch, I. (2019). Neural MMO: A Massively Multiagent Game Environment for Training and Evaluating Intelligent Agents.
Whiteson, S., Tanner, B., Taylor, M. E. ., & Stone, P. (2011). Protecting Against Evaluation Overfitting in Empirical Reinforcement Learning. In {IEEE} Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL).

IEEE TCDS Table of Contents

Volume 10, Issue 4, December 2018

Guest Editorial A Sense of Interaction in Humans and Robots: From Visual Perception to Social Cognition

A. Sciutti and N. Noceti

Human ability to interact with one another is substantially strengthened by vision, with several visual processes tuned to support prosocial behaviors since early infancy. A key challenge of robotics research is to provide artificial agents with similar advanced visual perception skills, with the ultimate goal of designing machines able to recognize and interpret both explicit and implicit communication cues embedded in human behaviors. This special issue addresses this challenge, with a focus both on understanding human perception supporting interaction abilities and on the implementation perspective, considering new algorithms and modeling efforts brought forward to improve current robotics. This multidisciplinary effort aims to bring innovations not only in human–machine interaction but also in domains such as developmental psychology and cognitive rehabilitation.

FFAB—The Form Function Attribution Bias in Human–Robot Interaction

K. S. Haring, K. Watanabe, M. Velonaki, C. C. Tossell and V. Finomore

People seem to miscalibrate their expectations and interactions with a robot. When it comes to robot design, the anthropomorphism level of the robot form (appearance) has become an increasingly important variable to consider. It is argued here that people base their expectations and perceptions of a robot on its form and attribute functions which do not necessarily mirror the true functions of the robot. The term form function attribution bias (FFAB) refers to the cognitive bias which occurs when people are prone to perceptual errors, leading to a biased interpretation of a robot's functionality. We argue that rather than objectively perceiving the robot's functionalities, people take a cognitive shortcut using the information available to them through visual perception. FFAB intends to outline the implications the design of a robot has on the human predisposition to interact socially with robots. In this theoretical review, we examined the results of several studies suggesting an FFAB. We outline future directions of experimental paradigms and robot design implications.

The Perception of Emotion in Artificial Agents

R. Hortensius, F. Hekele and E. S. Cross

Given recent technological developments in robotics, artificial intelligence, and virtual reality, it is perhaps unsurprising that the arrival of emotionally expressive and reactive artificial agents is imminent. However, if such agents are to become integrated into our social milieu, it is imperative to establish an understanding of whether and how humans perceive emotion in artificial agents. In this review, we incorporate recent findings from social robotics, virtual reality, psychology, and neuroscience to examine how people recognize and respond to emotions displayed by artificial agents. First, we review how people perceive emotions expressed by an artificial agent, such as facial and bodily expressions. Second, we evaluate the similarities and differences in the consequences of perceived emotions in artificial compared to human agents. Besides accurately recognizing the emotional state of an artificial agent, it is critical to understand how humans respond to those emotions. Does interacting with an angry robot induce the same responses in people as interacting with an angry person? Similarly, does watching a robot rejoice when it wins a game elicit similar feelings of elation in the human observer? Here, we provide an overview of the current state of emotion expression and perception during interactions with artificial agents, as well as a clear articulation of the challenges and guiding principles to be addressed as we move ever closer to truly emotional artificial agents.

Deep Construction of an Affective Latent Space via Multimodal Enactment

G. Boccignone, D. Conte, V. Cuculo, A. D'Amelio, G. Grossi and R. Lanzarotti

We draw on a simulationist approach to the analysis of facially displayed emotions, e.g., in the course of a face-to-face interaction between an expresser and an observer. At the heart of such perspective lies the enactment of the perceived emotion in the observer. We propose a novel probabilistic framework based on a deep latent representation of a continuous affect space, which can be exploited for both the estimation and the enactment of affective states in a multimodal space (visible facial expressions and physiological signals). The rationale behind the approach lies in the large body of evidence from affective neuroscience showing that when we observe emotional facial expressions, we react with congruent facial mimicry. Further, in more complex situations, affect understanding is likely to rely on a comprehensive

representation grounding the reconstruction of the state of the body associated with the displayed emotion. We show that our approach can address such problems in a unified and principled perspective, thus avoiding ad hoc heuristics while minimizing learning efforts.

Robot Fast Adaptation to Changes in Human Engagement During Simulated Dynamic Social Interaction With Active Exploration in Parameterized Reinforcement Learning

M. Khamassi, G. Velentzas, T. Tsitsimis and C. Tzafestas

Dynamic uncontrolled human-robot interactions (HRIs) require robots to be able to adapt to changes in the human's behavior and intentions. Among relevant signals, nonverbal cues such as the human's gaze can provide the robot with important information about the human's current engagement in the task, and whether the robot should continue its current behavior or not. However, robot reinforcement learning (RL) abilities to adapt to these nonverbal cues are still underdeveloped. Here, we propose an active exploration algorithm for RL during HRI where the reward function is the weighted sum of the human's current engagement and variations of this engagement. We use a parameterized action space where a meta-learning algorithm is applied to simultaneously tune the exploration in discrete action space (e.g., moving an object) and in the space of continuous characteristics of movement (e.g., velocity, direction, strength, and expressivity). We first show that this algorithm reaches state-of-the-art performance in the nonstationary multiarmed bandit paradigm. We then apply it to a simulated HRI task, and show that it outperforms continuous parameterized RL with either passive or active exploration based on different existing methods. We finally test the performance in a more realistic test of the same HRI task, where a practical approach is followed to estimate human engagement through visual cues of the head pose. The algorithm can detect and adapt to perturbations in human engagement with different durations. Altogether, these results suggest a novel efficient and robust framework for robot learning during dynamic HRI scenarios.

Features and Classification Schemes for View-Invariant and Real-Time Human Action Recognition

S. A. W. Talha, M. Hammouche, E. Ghorbel, A. Fleury and S. Ambellouis

Human action recognition (HAR) is largely used in the field of ambient assisted living to create an interaction between humans and computers. In these applications, it cannot be asked for people to act nonnaturally. The algorithm has to adapt and the interaction has to be as quick as possible to make it fluent. To improve the existing algorithms with regards to these points, we propose a novel method based on skeleton information provided by RGB-D cameras. This approach is able to carry out early action recognition and is more robust to viewpoint variability. To reach this goal, a new descriptor called body directional velocity is proposed and a real-time classification is performed. Experimental results on four benchmarks show that our method competes with various skeleton-based HAR algorithms. We also show the suitability of our method for early recognition of human actions.

Sensorimotor Communication for Humans and Robots: Improving Interactive Skills by Sending Coordination Signals

F. Donnarumma, H. Dindo and G. Pezzulo

During joint actions, humans continuously exchange coordination signals and use nonverbal, sensorimotor forms of communication. Here we discuss a specific example of sensorimotor communication-"signaling"-which consists in the intentional modification of one's own action plan (e.g., a plan for reaching a glass of wine) to make it more predictable or discriminable from alternative action plans that are contextually plausible (e.g., a plan for reaching another glass on the same table). We first review the existing evidence on signaling in human-human interactions, discussing under which conditions humans use signaling. Successively, we distill these insights into a computational theory of signaling during online interactions. Central to our approach are the following ideas: 1) signaling endows pragmatic plans with communicative goals; 2) signaling can be understood within a cost-benefit scheme, balancing the costs for the signaling agent against its benefits for interaction success; and 3) signaling may be part of an interactive strategy that optimizes success when joint goals are uncertain. Finally, we exemplify the benefits of signaling in a series of simulations and discuss how endowing robots with signaling abilities can increase the quality of human-robot interactions by making their behavior more predictable and "legible" for humans.

An Incremental Self-Organizing Architecture for Sensorimotor Learning and Prediction

L. Mici, G. I. Parisi and S. Wermter

During visuomotor tasks, robots must compensate for temporal delays inherent in their sensorimotor processing systems. Delay compensation becomes crucial in a dynamic environment where the visual

input is constantly changing, e.g., during the interaction with a human demonstrator. For this purpose, the robot must be equipped with a prediction mechanism for using the acquired perceptual experience to estimate possible future motor commands. In this paper, we present a novel neural network architecture that learns prototypical visuomotor representations and provides reliable predictions on the basis of the visual input. These predictions are used to compensate for the delayed motor behavior in an online manner. We investigate the performance of our method with a set of experiments comprising a humanoid robot that has to learn and generate visually perceived arm motion trajectories. We evaluate the accuracy in terms of mean prediction error and analyze the response of the network to novel movement demonstrations. Additionally, we report experiments with incomplete data sequences, showing the robustness of the proposed architecture in the case of a noisy visual input.

Assessing Social Competence in Visually Impaired People and Proposing an Interventional Program in Visually Impaired Children

Giulia Cappagli, Sara Finocchietti , Gabriel Baud-Bovy, Leonardo Badino, Alessandro D'Ausilio, Elena Cocchi, and Monica Gori

Visually impaired children and adults have difficulties in engaging in positive social interactions. This paper assesses social competence in sighted and visually impaired people and to propose a novel interventional strategy in visually impaired children. We designed a task that assesses the ability to initiate and sustain an interaction with the experimenter while performing free hand movements using a sonorous feedback on the experimenter's wrist. Both participant and experimenter kinematic data were recorded with a motion capture system. The level of social interaction between participant and experimenter has been computed through objective measurements based on Granger causality analysis applied to the participant arm kinematics. The interventional program followed by the visually impaired children lasted 12 weeks and consisted in a series of spatial and social games performed with the use of a sonorous bracelet which provides an auditory feedback of body actions in space. Visually impaired individuals present a poorer communication flow with the experimenter than sighted people, which indicates a less efficient social interaction. The amount of communication between the two agents resulted in a significant improvement after the interventional program. Thus, a specific intervention, based on the substitution of visual with auditory feedback of body actions, can enhance social inclusion for the blind population.

Toward Socially Aware Person-Following Robots

S. S. Honig, T. Oron-Gilad, H. Zaichyk, V. Sarne-Fleischmann, S. Olatunji and Y. Edan

Significant research and development has been invested in technical issues related to person following. However, a systematic approach for designing robotic person-following behavior that maintains appropriate social conventions across contexts has not yet been developed. To understand why this may be the case, an in-depth literature review of 221 articles on person-following robots was performed, from which 107 are referenced. From these papers, six relevant topics were identified that shed light on the types of social interactions that have been studied in person-following scenarios: 1) applications; 2) robotic systems; 3) environments; 4) following strategies; 5) human-robot communication; and 6) evaluation methods. Gaps in the existing research on person-following robots were identified, mainly in addressing social interaction and user needs, noting that only 25 articles reported proper user studies. Human-related, robot-related, task-related, and environment-related factors that are likely to influence people's spatial preferences and expectations of a robot's person-following behavior are then discussed. To guide the design of socially aware person following robots, a user-needs layered design framework that combines the four factor categories is proposed. The framework provides a systematic way to incorporate social considerations in the design of person-following robots. Finally, framework limitations and future challenges in the field are presented and discussed.

Extending Human–Robot Relationships Based in Music With Virtual Presence

L. McCallum and P. W. McOwan

Social relationships between humans and robots require both long term engagement and a feeling of believability or social presence toward the robot. It is our contention that music can provide the extended engagement that other open-ended interaction studies have failed to do, also, that in combination with the engaging musical interaction, the addition of simulated social behaviors is necessary to trigger this sense of believability or social presence. Building on previous studies with our robot drummer Mortimer that show including social behaviors can increase engagement and social presence, we present the results of a longitudinal study investigating the effect of extending weekly collocated musical improvisation sessions by making Mortimer an active member of the participant's virtual social network. Although, we found the effects of extending the relationship into the virtual world were less pronounced than results we have previously found by adding social modalities to human-robot musical interaction, interesting questions are raised about the interpretation of our automated behavioral metrics across different contexts. Further, we found repeated results of increasingly uninteruppted playing and notable differences in responses to online posts by Mortimer and posts by participant's human friends.

Action Augmented Real Virtuality: A Design for Presence

M. S. L. Khan, A. Halawani, S. ur Réhman and H. Li

This paper addresses the important question of how to design a video teleconferencing setup to increase the experience of spatial and social presence. Traditional video teleconferencing setups are lacking in presenting the nonverbal behaviors that humans express in face-to-face communication, which results in decrease in presence-experience. In order to address this issue, we first present a conceptual framework of presence for video teleconferencing. We introduce a modern presence concept called real virtuality and propose a new way of achieving this based on body or artifact actions to increase the feeling of presence, and we named this concept presence through actions. Using this new concept, we present the design of a novel action-augmented real virtuality prototype that considers the challenges related to the design of an action prototype, action embodiment, and face representation. Our action prototype is a telepresence mechatronic robot (TEBoT), and action embodiment is through a head-mounted display (HMD). The face representation solves the problem of face occlusion introduced by the HMD. The novel combination of HMD, TEBoT, and face representation algorithm has been tested in a real video teleconferencing scenario for its ability to solve the challenges related to spatial and social presence. We have performed a user study where the invited participants were requested to experience our novel setup and to compare it with a traditional video teleconferencing setup. The results show that the action capabilities not only increase the feeling of spatial presence but also increase the feeling of social presence of a remote person among local collaborators.

Uncoupling Between Multisensory Temporal Function and Nonverbal Turn-Taking in Autism Spectrum Disorder

J. Noel, M. A. De Niear, N. S. Lazzara and M. T. Wallace

The integration of information across distinct modalities enhances perceptual abilities. An ecologically important role of multisensory integration is in scaffolding verbal communication, which relies upon the precise temporal integration of auditory and visual cues. However, the role of (multi)sensory function in supporting another important aspect of communication, namely, nonverbal communication, is unknown. Here, individuals with autism spectrum disorder (ASD) and a group of typically developing (TD) participants performed a simultaneity judgment task to index their audiovisual temporal acuity for speech stimuli. Further, under a naturalistic scenario, nonverbal synchrony between the participant and a naïve experimenter was measured. Automated motion analysis was performed to quantify movements of different body-parts. Results demonstrate a wider window of audiovisual temporal integration for ASD participants in comparison to their TD counterparts. Moreover, ASD individuals performed less complex movements and demonstrated less nonverbal synchrony during the interactive exchange. Lastly, multisensory temporal acuity significantly predicted the synchrony in hand and head movements between TD participants and the experimenter, but not between the ASD participants and the experimenter. Taken together, the results suggest an important role for multisensory perceptual abilities in shaping nonverbal communication between dyads and highlight the important role of perceptual systems in supporting social interactive skills.

Toward Improved Child–Robot Interaction by Understanding Eye Movements

K. S. Lohan, E. Sheppard, G. Little and G. Rajendran

Globally, 1 in 160 children has an autism spectrum disorder (ASD). Problems with joint attention (JA) are core features of ASDs. Here, we investigate how typically developing (TD) children and children with ASD initiate JA with a gaze contingent avatar. Thirty-one participants with ASD and 33 TD matched controls directed an avatar's gaze to a series of referent images. Observing pupil diameter and gaze location data, we explore how distinguishing the two groups as well as their different eye-movement behaviors could be used to improve child-robot interaction. With a sequence to sequence neural network we distinguish if a child is TD or has an ASD, then using K-means clustering, we group pupil diameters and gaze locations independently to determine the child's attention level as well as to refine the classification process. Using these metrics, we could trigger appropriate responses from the robot to increase the level of attention from the child toward the robot. Results show significant differences between the eye behaviors of individuals with ASDs and those without. Further to this, we achieve a 79.76% classification accuracy when using pupil diameter data to distinguish the two groups.

Study of Mechanisms of Social Interaction Stimulation in Autism Spectrum Disorder by Assisted Humanoid Robot

M. Del Coco, M. Leo, P. Carcagnì, F. Famà, L. Spadaro, L. Ruta, G. Pioggia, and C. Distante

Information and communication technologies (ICTs) have been proved to have a great impact in enhancing social, communicative, and language development in children with autism spectrum disorders (ASDs) as demonstrated by plenty of effective technological tools reported in the literature for diagnosis, assessment, and treatment of such neurological diseases. On the contrary, there are very few works exploiting ICT to study the mechanisms that trigger the behavioral patterns during the specialized sessions of

treatment focused on social interaction stimulation. From the study of the literature it emerges that the behavioral outcomes are qualitatively evaluated by the therapists making this way impossible to assess, in a consistent manner, the worth of the supplied ASD treatments that should be based on quantitative metric not available for this purpose yet. Moreover, the rare attempts to use a methodological approach are limited to the study of one (of at least a couple) of the several behavioral cues involved. In order to fill this gap, in this paper a technological framework able to analyze and integrate multiple visual cues in order to capture the behavioral trend along an ASD treatment is introduced. It is based on an algorithmic pipeline involving face detection, landmark extraction, gaze estimation, head pose estimation and facial expression recognition and it has been used to detect behavioral features during the interaction among different children, affected by ASD, and a humanoid robot. Experimental results demonstrated the superiority of the proposed framework in the specific application context with respect to leading approaches in the literature, providing a reliable pathway to automatically build a quantitative report that could help therapists to better achieve either ASD diagnosis or assessment tasks.

DAC-h3: A Proactive Robot Cognitive Architecture to Acquire and Express Knowledge About the World and the Self

C. Moulin-Frier, T. Fischer, M. Petit, G. Pointeau, J. Puigbo, U. Pattacini, S. C. Low, D. Camilleri, P. Nguyen, M. Hoffmann, H. J. Chang, M. Zambelli, A. Mealier, A. Damianou, G. Metta, T. J. Prescott, Y. Demiris, P. F. Dominey, P. F. M. J. Verschure

This paper introduces a cognitive architecture for a humanoid robot to engage in a proactive, mixed-initiative exploration and manipulation of its environment, where the initiative can originate from both human and robot. The framework, based on a biologically grounded theory of the brain and mind, integrates a reactive interaction engine, a number of state-of-the-art perceptual and motor learning algorithms, as well as planning abilities and an autobiographical memory. The architecture as a whole drives the robot behavior to solve the symbol grounding problem, acquire language capabilities, execute goal-oriented behavior, and express a verbal narrative of its own experience in the world. We validate our approach in human-robot interaction experiments with the iCub humanoid robot, showing that the proposed cognitive architecture can be applied in real time within a realistic scenario and that it can be used with naive users.

Comparison Studies on Active Cross-Situational Object-Word Learning Using Non-Negative Matrix Factorization and Latent Dirichlet Allocation

Y. Chen, J. Bordes and D. Filliat

Future intelligent robots are expected to be able to adapt continuously to their environment. For this purpose, recognizing new objects and learning new words through interactive learning with humans is fundamental. Such setup results in ambiguous teaching data which humans have been shown to address using cross-situational learning, i.e., by analyzing common factors between multiple learning situations. Moreover, they have been shown to be more efficient when actively choosing the learning samples, e.g., which object they want to learn. Implementing such abilities on robots can be performed by latent-topic learning models such as non-negative matrix factorization or latent Dirichlet allocation. These cross-situational learning methods tackle referential and linguistic ambiguities, and can be associated with active learning strategies. We propose two such methods: 1) the maximum reconstruction error-based selection and 2) confidence base exploration. We present extensive experiments using these two learning algorithms through a systematic analysis on the effects of these active learning strategies in contrast with random choice. In addition, we study the factors underlying the active learning by focusing on the use of sample repetition, one of the learning behaviors that have been shown to be important for humans.

The Necessary and Sufficient Conditions for Emergence in Systems Applied to Symbol Emergence in Robots

R. L. Sturdivant and E. K. P. Chong

A conceptual model for emergence with downward causation is developed. In addition, the necessary and sufficient conditions are identified for a phenomenon to be considered emergent in a complex system. It is then applied to symbol emergence in robots. This paper is motivated by the usefulness of emergence to explain a wide variety of phenomena in systems, and cognition in natural and artificial creatures. Downward causation is shown to be a critical requirement for potentially emergent phenomena to be considered actually emergent. Models of emergence with and without downward causation are described and how weak emergence can include downward causation. A process flow is developed for distinguishing emergence from nonemergence based upon the application of reductionism and detection of downward causation. Examples are shown for applying the necessary and sufficient conditions to filter out actually emergent phenomena from nonemergent ones. Finally, this approach for detecting emergence is applied to complex projects and symbol emergence in robots.

Ensemble-of-Concept Models for Unsupervised Formation of Multiple Categories

T. Nakamura and T. Nagai

Recent studies have shown that robots can form concepts and understand the meanings of words through inference. The key idea underlying these studies is the "multimodal categorization" of a robot's experiences. Despite the success in the formation of concepts by robots, a major drawback of previous studies stems from the fact that they have been mainly focused on object concepts. Obviously, human concepts are limited not only to object concepts but also to other kinds such as those connected to the tactile sense and color. In this paper, we propose a novel model called the ensemble-of-concept models (EoCMs) to form various kinds of concepts. In EoCMs, we introduce weights that represent the strength connecting modalities and concepts. By changing these weights, many concepts that are connected to particular modalities can be formed; however, meaningless concepts for humans are included in these concepts. To communicate with humans, robots are required to form meaningful concepts for us. Therefore, we utilize utterances taught by human users as the robot observes objects. The robot connects words included in the teaching utterances with formed concepts and selects meaningful concepts to communicate with users. The experimental results show that the robot can form not only object concepts but also others such as color-related concepts and haptic concepts. Furthermore, using word2vec, we compare the meanings of the words acquired by the robot in connecting them to the concepts formed.

Recognition of Visually Perceived Compositional Human Actions by Multiple Spatio-Temporal Scales Recurrent Neural Networks

H. Lee, M. Jung and J. Tani

We investigate a deep learning model for action recognition that simultaneously extracts spatio-temporal information from a raw RGB input data. The proposed multiple spatio-temporal scales recurrent neural network (MSTRNN) model is derived by combining multiple timescale recurrent dynamics with a conventional convolutional neural network model. The architecture of the proposed model imposes both spatial and temporal constraints simultaneously on its neural activities. The constraints vary, with multiple scales in different layers. As suggested by the principle of upward and downward causation, it is assumed that the network can develop a functional hierarchy using its constraints during training. To evaluate and observe the characteristics of the proposed model, we use three human action datasets consisting of different primitive actions and different compositionality levels. The performance capabilities of the MSTRNN model on these datasets are compared with those of other representative deep learning models used in the field. The results show that the MSTRNN outperforms baseline models while using fewer parameters. The characteristics of the proposed model are observed by analyzing its internal representation properties. The analysis clarifies how the spatio-temporal constraints of the MSTRNN model aid in how it extracts critical spatio-temporal information relevant to its given tasks.

Online Covariate Shift Detection-Based Adaptive Brain–Computer Interface to Trigger Hand Exoskeleton Feedback for Neuro-Rehabilitation

A. Chowdhury, H. Raza, Y. K. Meena, A. Dutta and G. Prasad

A major issue in electroencephalogram (EEG)-based brain-computer interfaces (BCIs) is the intrinsic nonstationarities in the brain waves, which may degrade the performance of the classifier, while transitioning from calibration to feedback generation phase. The nonstationary nature of the EEG data may cause its input probability distribution to vary over time, which often appear as a covariate shift. To adapt to the covariate shift, we had proposed an adaptive learning method in our previous work and tested it on offline standard datasets. This paper presents an online BCI system using previously developed covariate shift detection (CSD)-based adaptive classifier to discriminate between mental tasks and generate neurofeedback in the form of visual and exoskeleton motion. The CSD test helps prevent unnecessary retraining of the classifier. The feasibility of the developed online-BCI system was first tested on ten healthy individuals, and then on ten stroke patients having hand disability. A comparison of the proposed online CSD-based adaptive classifier with conventional nonadaptive classifier has shown a significantly ($p <; 0.01$) higher classification accuracy in both the cases of healthy and patient groups. The results demonstrate that the online CSD-based adaptive BCI system is superior to the nonadaptive BCI system and it is feasible to be used for actuating hand exoskeleton for the stroke-rehabilitation applications.

A Self-Verifying Cognitive Architecture for Robust Bootstrapping of Sensory-Motor Skills via Multipurpose Predictors

E. Wieser and G. Cheng

The autonomous acquisition of sensory-motor skills along multiple developmental stages is one of the current challenges in robotics. To this end, we propose a new developmental cognitive architecture that combines multipurpose predictors and principles of self-verification for the robust bootstrapping of sensory-motor skills. Our architecture operates with loops formed by both mental simulation of sensory-motor sequences and their subsequent physical trial on a robot. During these loops, verification algorithms monitor the predicted and the physically observed sensory-motor data. Multiple types of predictors are acquired through several developmental stages. As a result, the architecture can select

and plan actions, adapt to various robot platforms by adjusting proprioceptive feedback, predict the risk of self-collision, learn from a previous interaction stage by validating and extracting sensory-motor data for training the predictor of a subsequent stage, and finally acquire an internal representation for evaluating the performance of its predictors. These cognitive capabilities in turn realize the bootstrapping of early hand-eye coordination and its improvement. We validate the cognitive capabilities experimentally and, in particular, show an improvement of reaching as an example skill.

An Experimental Study of Embodied Interaction and Human Perception of Social Presence for Interactive Robots in Public Settings

D. C. Herath, E. Jochum and E. Vlachos

The human perception of robots as social depends on many factors, including those that do not necessarily pertain to a robot's cognitive functioning. Experience design (ED) offers a useful framework for evaluating when participants interact with robots as products or tools and when they regard them as social actors. This paper describes a between-participants experiment conducted at a science museum, where visitors were invited to play a game of noughts and crosses with a Baxter robot. The goal is to foster meaningful interactions that promote engagement between the human and the robot in a museum context. Using an ED framework, we tested the robot in three different conditions to better understand which factors contribute to the perception of robots as social. The experiment also outlines best practices for conducting human-robot interaction research in museum exhibitions. Results from the study indicate that perceived social presence can be evaluated using a combination of human robot interaction and ED methods that measure co-presence and co-experience.

Volume 11, Issue 1, March 2019

Toward Brain-Inspired Learning With the Neuromorphic Snake-Like Robot and the Neurorobotic Platform

Guang Chen, Zhenshan Bing, Florian Röhrbein, Jörg Conradt, Kai Huang, Long Cheng, Zhuangyi Jiang, Alois Knoll

Neurorobotic mimics the structural and functional principles of living creature systems. Modeling a single system by robotic hardware and software has existed for decades. However, an integrated toolset studying the interaction of all systems has not been demonstrated yet. We present a hybrid neuromorphic computing paradigm to bridge this gap by combining the neurorobotics platform (NRP) with the neuromorphic snake-like robot (NeuroSnake). This paradigm encompasses the virtual models, neuromorphic sensing and computing capabilities, and physical bio-inspired bodies, with which an experimenter can design and execute both in-silico and in-vivo robotic experimentation easily. The NRP is a public Web-based platform for easily testing brain models with virtual bodies and environments. The NeuroSnake is a bio-inspired robot equipped with a silico-retina sensor and neuromorphic computer for power-efficiency applications. We illustrate the efficiencies of our paradigm with an easy designing of a visual pursuit experiment in the NRP. We study two automatic behavior learning tasks which are further integrated into a complex task of semi-autonomous pole climbing. The result shows that robots could build new learning rules in a less explicit manner inspired by living creatures. Our method gives an alternative way to efficiently develop complex behavior control of the ro As spiking neural network is a bio-inspired neural network and the NeuroSnake robot is equipped with a spike-based silicon retina camera, the control system can be easily implemented via spiking neurons simulated on neuromorphic hardware, such as SpiNNaker.bot.

Brain-Inspired Cognitive Model With Attention for Self-Driving Cars

S. Chen, S. Zhang, J. Shang, B. Chen and N. Zheng

The perception-driven approach and end-to-end system are two major vision-based frameworks for self-driving cars. However, it is difficult to introduce attention and historical information into the autonomous driving process, which are essential for achieving human-like driving in these two methods. In this paper, we propose a novel model for self-driving cars called the brain-inspired cognitive model with attention. This model comprises three parts: 1) a convolutional neural network for simulating the human visual cortex; 2) a cognitive map to describe the relationships between objects in a complex traffic scene; and 3) a recurrent neural network, which is combined with the real-time updated cognitive map to implement the attention mechanism and long-short term memory. An advantage of our model is that it can accurately solve three tasks simultaneously: 1) detecting the free space and boundaries for the current and adjacent lanes; 2) estimating the distances to obstacles and vehicle attitude; and 3) learning the driving behavior and decision-making process of a human driver. Importantly, the proposed model can accept external navigation instructions during an end-to-end driving process. To evaluate the

model, we built a large-scale road-vehicle dataset containing over 40 000 labeled road images captured by three cameras placed on our self-driving car. Moreover, human driving activities and vehicle states were recorded at the same time.

Combining Model-Based Q -Learning With Structural Knowledge Transfer for Robot Skill Learning

Z. Deng, H. Guan, R. Huang, H. Liang, L. Zhang and J. Zhang

Learning skills autonomously is a particularly important ability for an autonomous robot. A promising approach is reinforcement learning (RL) where agents learn policy through interaction with its environment. One problem of RL algorithm is how to tradeoff the exploration and exploitation. Moreover, multiple tasks also make a great challenge to robot learning. In this paper, to enhance the performance of RL, a novel learning framework integrating RL with knowledge transfer is proposed. Three basic components are included: 1) probability policy reuse; 2) dynamic model learning; and 3) model-based ${Q}$ -learning. In this framework, the prelearned skills are leveraged for policy reuse and dynamic learning. In model-based ${Q}$ -learning, the Gaussian process regression is used to approximate the ${Q}$ -value function so as to suit for robot control. The prior knowledge retrieved from knowledge transfer is integrated into the model-based ${Q}$ -learning to reduce the needed learning time. Finally, a human-robot handover experiment is performed to evaluate the learning performance of this learning framework. Experiment results show that fewer exploration is needed to obtain a high expected reward, due to the prior knowledge obtained from knowledge transfer.

Human Behavior-Based Target Tracking With an Omni-Directional Thermal Camera

E. Benli, Y. Motai and J. Rogers

We investigate human behavior-based target tracking from omni-directional (O-D) thermal images for intelligent perception in unmanned systems. Current target tracking approaches are primarily focused on perspective visual and infrared (IR) band, as well as O-D visual band tracking. The target tracking from O-D images and the use of O-D thermal vision have not been adequately addressed. Thermal O-D images provide a number of advantages over other passive sensor modalities such as illumination invariance, wide field-of-view, ease of identifying heat-emitting objects, and long term tracking without interruption. Unfortunately, thermal O-D sensors have not yet been widely used due to the following disadvantages: low resolution, low frame rates, high cost, sensor noise, and an increase in tracking time. This paper outlines a spectrum of approaches which mitigate these disadvantages to enable an O-D thermal IR camera equipped with a mobile robot to track a human in a variety of environments and conditions. The curve matched Kalman filter is used for tracking a human target based on the behavioral movement of the human and maximum a posteriori (MAP)-based estimation is extended for the human tracking as long term which provides a faster prediction. The benefits to using our MAP-based method are decreasing the prediction time of a target's position and increasing the accuracy of prediction of the next target position based on the target's previous behavior while increasing the tracking view and lighting conditions via the view from O-D IR camera.

Brain-Actuated Control of Dual-Arm Robot Manipulation With Relative Motion

Z. Li, W. Yuan, S. Zhao, Z. Yu, Y. Kang and C. L. P. Chen

This paper describes the brain-actuated control of a dual-arm robot performing bimanual relative motion manipulation tasks, through the adoption of relative Jacobian matrix, wherein the dual-arm robot can be considered as a single manipulator and the movements of the end effectors can be calculated by the relative motion. An online brain–machine interface (BMI) system based on multichannel steady-state visual evoked potentials was developed, and it was able to perform band-pass filtering and visual stimuli classification using support vector machines. Considering the relative motion in a constrained plane, the asymmetric bimanual manipulation can be transformed into 2-D control tasks through polar coordinate transformation such that the end effectors can achieve smooth direction-and-distance motion in the arbitrary position of the operational space. Moreover, the kinematics redundancy scheme, using online neuro-dynamics optimization, was developed for joint velocity optimization subject to physical constraints. Five individuals participated in the experiments and successfully fulfilled the given manipulation task.

A Computational Model of Neuroreceptor-Dependent Plasticity (NRDP) Based on Spiking Neural Networks

J. I. Espinosa-Ramos, E. Capecci and N. Kasabov

Activity-dependent plasticity has attracted the interest of researchers for years in the domain of computational neuroscience, as the modification of synaptic efficacy occurs as a result of complex biochemical mechanisms that take place at a cellular level. In this paper, we introduce a phenomenological model implemented as an unsupervised learning rule for spiking neural networks (SNNs) based on the cross-talk

between glutamatergic and ${\gamma }$ -aminobutyric acid (GABA)ergic neuroreceptors: ${N}$ -methyl-${D}$ -aspartate receptor, ${\alpha }$ -amino-3-hydroxy-5-methyl-4-izoxazole-propionic acid receptors, GABAA, and GABAB. The proposed neuroreceptor-dependent plasticity (NRDP) model is implemented and demonstrated in an SNN environment, NeuCube, for modeling electroencephalography data. We show that the NRDP model can reproduce the generic spike-timing dependent plasticity behavior in an SNN. In addition, this can be used to simulate changes in excitatory/inhibitory balance in an SNN by altering neuroreceptors activity. More specifically, by varying the parameters that affect neuroreceptors activation, we can study how these changes would affect the learning and memory ability of a subject. In a therapeutic context, this makes it a promising tool for studying the regulatory mechanisms, where neuroreceptors cross-talk plays a crucial role. This can lead to new ways of early detection of neurological disorders and for better targeting drug treatments.

Complete Coverage Autonomous Underwater Vehicles Path Planning Based on Glasius Bio-Inspired Neural Network Algorithm for Discrete and Centralized Programming

B. Sun, D. Zhu, C. Tian and C. Luo

For the complete coverage path planning of autonomous underwater vehicles (AUVs), a new strategy with Glasius bio-inspired neural network (GBNN) algorithm with discrete and centralized programming is proposed. The basic modeling for multi-AUVs complete coverage problem based on grid map and neural network is discussed first. Then, the design for single AUV complete coverage is introduced based on GBNN algorithm which is a new developed tool with small amount of calculation and high efficiency. In order to solve the difficulty of single AUV full coverage task of large water range, the multi-AUV full coverage discrete and centralized programming is proposed based on GBNN algorithm. The simulation experiment is conducted to confirm that through the proposed algorithm, multi-AUVs can plan reasonable and collision-free coverage path and reach full coverage on the same task area with division of labor and cooperation.

Domain Adaptation Techniques for EEG-Based Emotion Recognition: A Comparative Study on Two Public Datasets

Z. Lan, O. Sourina, L. Wang, R. Scherer and G. R. Müller-Putz

Affective brain–computer interface (aBCI) introduces personal affective factors to human–computer interaction. The state-of-the-art aBCI tailors its classifier to each individual user to achieve accurate emotion classification. A subject-independent classifier that is trained on pooled data from multiple subjects generally leads to inferior accuracy, due to the fact that electroencephalography patterns vary from subject to subject. Transfer learning or domain adaptation techniques have been leveraged to tackle this problem. Existing studies have reported successful applications of domain adaptation techniques on SEED dataset. However, little is known about the effectiveness of the domain adaptation techniques on other affective datasets or in a cross-dataset application. In this paper, we focus on a comparative study on several state-of-the-art domain adaptation techniques on two datasets: 1) DEAP and 2) SEED. We demonstrate that domain adaptation techniques can improve the classification accuracy on both datasets, but not so effective on DEAP as on SEED. Then, we explore the efficacy of domain adaptation in a cross-dataset setting when the data are collected under different environments using different devices and experimental protocols. Here, we propose to apply domain adaptation to reduce the intersubject variance as well as technical discrepancies between datasets, and then train a subject-independent classifier on one dataset and test on the other. Experiment results show that using domain adaptation technique in a transductive adaptation setting can improve the accuracy significantly by 7.25%–13.40% compared to the baseline accuracy where no domain adaptation technique is used.

A Novel Brain Decoding Method: A Correlation Network Framework for Revealing Brain Connections

S. Yu, N. Zheng, Y. Ma, H. Wu and B. Chen

Brain decoding is a hot spot in cognitive science, which focuses on reconstructing perceptual images from brain activities. Analyzing the correlations of collected data from human brain activities and representing activity patterns are two key problems in brain decoding based on functional magnetic resonance imaging signals. However, existing correlation analysis methods mainly focus on the strength information of voxel, which reveals functional connectivity in the cerebral cortex. They tend to neglect the structural information that implies the intracortical or intrinsic connections; that is, structural connectivity. Hence, the effective connectivity inferred by these methods is relatively unilateral. Therefore, we propose in this paper a correlation network (CorrNet) framework that could be flexibly combined with diverse pattern representation models. In the CorrNet framework, the topological correlation is introduced to reveal structural information. Rich correlations can be obtained, which contribute to specifying the underlying effective connectivity. We also combine the CorrNet framework with a linear support vector machine and a dynamic evolving spike neuron network for pattern representation separately, thus provide a novel method for decoding cognitive activity patterns. Experimental results verify the reliability and robustness of our CorrNet framework, and demonstrate that the new method can achieve significant improvement

in brain decoding over comparable methods.

Canonical Correlation Analysis Regularization: An Effective Deep Multiview Learning Baseline for RGB-D Object Recognition

L. Tang, Z. Yang and K. Jia

Object recognition methods based on multimodal data, color plus depth (RGB-D), usually treat each modality separately in feature extraction, which neglects implicit relations between two views and preserves noise from any view to the final representation. To address these limitations, we propose a novel canonical correlation analysis (CCA)-based multiview convolutional neural network (CNNs) framework for RGB-D object representation. The RGB and depth streams process corresponding images, respectively, then are connected by CCA module leading to a common-correlated feature space. In addition, to embed CCA into deep CNNs in a supervised manner, two different schemes are explored. One considers CCA as a regularization (CCAR) term adding to the loss function. However, solving CCA optimization directly is neither computationally efficient nor compatible with the mini-batch-based stochastic optimization. Thus, we further propose an approximation method of CCAR, using the obtained CCA projection matrices to replace the weights of feature concatenation layer at regular intervals. Such a scheme enjoys benefits of full CCAR and is efficient by amortizing its cost over many training iterations. Experiments on benchmark RGB-D object recognition datasets have shown that the proposed methods outperform most existing methods using the very same of their network architectures.

Adaptive Drawing Behavior by Visuomotor Learning Using Recurrent Neural Networks

K. Sasaki and T. Ogata

Drawing is a medium that represents an idea as drawn lines, and drawing behavior requires complex cognitive abilities to process visual and motor information. One way to understand aspects of these abilities is constructing computational models that can replicate these abilities rather than explaining the phenomena by building plausible models by a top-down manner. In this paper, we proposed a supervised learning model that can be trained using examples of visuomotor sequences from drawings made by human. Additionally, we demonstrated that the proposed model has functions of: 1) associating motions to depict the given picture image and 2) adapting to drawing behavior to complete a given part of the drawing process. This dynamical model is implemented by recurrent neural networks that have images and motion as their input and output. Through experiments that involved learning human drawing sequences, the model was able to associate appropriate motions to achieve depiction targets while adapting to a given part of the drawing process. Furthermore, we demonstrate that including visual information in the model improved performance robustness against noisy lines in the input data.

Brain–Computer Interface-Based Stochastic Navigation and Control of a Semiautonomous Mobile Robot in Indoor Environments

Y. Yuan, W. Su, Z. Li and G. Shi

In this paper, a brain–computer interface (BCI)-based navigation and control strategy is developed for a mobile robot in indoor environments. It combines the simultaneous localization and mapping to achieve the navigation and positioning for a mobile robot in indoor environments, where the RGB landmarks are regarded as the environmental features learned by the FastSLAM algorithm. The online BCI, based on steady-state visually evoked potentials, exploits multivariate synchronization index classification algorithm to analyze the human electroencephalograph (EEG) signals so that the human intention can be recognized accurately, and then the EEG-based motion commands are produced for the mobile robot. Probability potential field approach based on the probability density function of 2-D normal distribution is connected with the brain signals to generate a collision-free trajectory for the mobile robot. The entire system is semiautonomous, since the robot's low level behaviors are autonomous and the stochastic navigation is executed by the BCI, and it is verified by the extensive experiments involving five volunteers. All the participants can successfully tele-operate the mobile robot, and the experimental results have verified the effectiveness of the proposed approach.

Volume 11, Issue 2, June 2019

Guest Editorial Neuro-Robotics Systems: Sensing, Cognition, Learning, and Control

Zhijun Li, Fei Chen, Antonio Bicchi, Yu Sun, Toshio Fukuda

Neuro-robotics systems (NRSs) are the most advanced research in the field of robotics, promoting the seamless exchange of knowledge between neuroscience and robotics. Recent breakthroughs in human brain neuroscience could be applied to robots to improve their perception, cognition, learning, and control abilities, which even help to develop a brain for robots. A challenge of robotics research is to break through technical bottlenecks by the findings in neuroscience and make robots move more flexibly, have more intelligent perception and control, and communicate with humans more naturally. To address the challenge, this special issue focuses on the latest advances in the area of NRS, particularly, the research concentrating on human-like sensing, multimodal information fusion, cognition, learning, and control technology. The development of these technologies aims to bring innovations into robotics and NRSs.

Combined Sensing, Cognition, Learning, and Control for Developing Future Neuro-Robotics Systems: A Survey

Junjun Li, Zhijun Li, Fei Chen, Antonio Bicchi, Yu Sun, Toshio Fukuda

Neuro-robotics systems (NRSs) is the current state-of-the-art research with the strategic alliance of neuroscience and robotics. It endows the next generation of robots with embodied intelligence to identify themselves and interact with humans and environments naturally. Therefore, it needs to study the interaction of recent breakthroughs in brain neuroscience, robotics, and artificial intelligence where smarter robots could be developed by employing neural mechanisms and understanding brain functions. Recently, more sophisticated neural mechanisms of perception, cognition, learning, and control have been decoded, which investigate how to define and develop the "brain" for future robots. In this paper, a comprehensive survey is summarized by recent achievements in neuro-robotics, and some potential directions for the development of future neuro-robotics are discussed.

Multimodal Human Hand Motion Sensing and Analysis—A Review

Yaxu Xue, Zhaojie Ju, Kui Xiang, Jing Chen, Honghai Liu

Human hand motion (HHM) analysis is an essential research topic in recent applications, especially for dexterous robot hand manipulation learning from human hand skills. It provides important information about the gestures, tactile, speed, and contact force, captured via multiple sensing technologies. This paper introduces a comprehensive survey of current hand motion sensing technologies and analysis approaches in recent emerging applications. First, the nature of HHMs is discussed in terms of simple motions, such as grasps and gestures, and complex motions, e.g., in-hand manipulations and regrasps; second, different techniques for hand motion sensing, including contact-based and noncontact-based approaches, are discussed with comparisons with their pros and cons; then, the state-of-theart analysis methods are introduced, with a particular focus on the multimodal hand motion sensing and analysis; finally, cuttingedge applications of hand motion analysis are reviewed, with further discussion on facing challenges and new future directions.

Active Visual-Tactile Cross-Modal Matching

Huaping Liu, Feng Wang, Fuchun Sun, Xinyu Zhang

Tactile and visual modalities frequently occur in cognitive robotics. Their matching problem is of highly interesting in many practical scenarios since it provides different properties about objects. In this paper, we investigate the active visual-tactile cross-modal matching problem which is formulated as retrieving the relevant sample in unlabeled gallery visual dataset in response to the tactile query sample. Such a problem exhibits a nontrivial challenge that there does not exist sample-to-sample pairing relation between tactile and visual modalities. To this end, we design a shared dictionary learning model which can simultaneously learn the projection subspace and the latent shared dictionary for the visual and tactile measurements. In addition, an optimization algorithm is developed to effectively solve the shared dictionary learning problem. Based on the obtained solution, the visual-tactile cross-modal matching algorithm can be easily developed. Finally, we perform experimental validations on the PHAC-2 datasets to show the effectiveness of the proposed visual-tactile cross-modal matching framework and method.

Fused Fuzzy Petri Nets: A Shared Control Method for Brain–Computer Interface Systems

Fuchun Sun, Wenchang Zhang, Jianhua Chen, Hang Wu, Chuanqi Tan, Weihua Su

It is hard to grasp objects based on brain-computer interface (BCI) by brain-actuated robot arm and hand

due to its high degree of freedom. Shared control strategy and hybrid BCI are research trends to solve this control problem of brainactuated discrete event system. We propose a new shared control method based on fused fuzzy Petri nets (PNs) for combining the robot automatic control (AC) and the brain-actuated control. This method takes the advantages of both fuzzy control and PNs such as easy modeling, robustness, and effectiveness. Both MATLAB simulation test and Barrett robot hand practical experiments show that the proposed method performs much better than AC or BCI control independently. In the online BCI practical experiment, the user successfully control the Barrett robot hand to grasp object avoiding obstacle in whole ten random scenes by our shared control method and hybrid BCI. Compared with BCI control, the user needs not to synchronously work according to the specific paradigm in the whole process. Meanwhile, our method improves safety and robustness by comparing with AC.

Bio-Inspired Equilibrium Point Control Scheme for Quadrupedal Locomotion

Yapeng Shi, Pengfei Wang, Xin Wang, Fusheng Zha, Zhenyu Jiang, Wei Guo, Mantian Li

The goal of this paper is to present a force control scheme for quadrupedal locomotion by adopting observations of biological motor behavior. Specifically, based on the equilibrium point (EP) hypothesis, we set up a bio-inspired EP controller in Cartesian space. The proposed EP controller modifies the EP trajectories appropriately over time from two perspectives, which can ensure stable interactions and system equilibrium. One perspective is directly compensating for the posture angle error based on torso posture compensation. The other is based on the foot force tracking algorithm and admittance model. The main contribution of this paper is to show how the EP controller realizes dynamic balance with no input about inertial parameter identification for the robot or terrain information estimation. Overall, the EP control scheme is simplified for stability problems and is easy to use in practice. Finally, we carried out a series of simulations and experiments to evaluate the effectiveness of the EP control algorithm. The results demonstrate that the proposed controller may improve dynamic stability and realize compliance performance.

Episodic Memory Multimodal Learning for Robot Sensorimotor Map Building and Navigation

Wei Hong Chin, Yuichiro Toda, Naoyuki Kubota, Chu Kiong Loo, Manjeevan Seera

In this paper, an unsupervised learning model of episodic memory is proposed. The proposed model, enhanced episodic memory adaptive resonance theory (EEM-ART), categorizes and encodes experiences of a robot to the environment and generates a cognitive map. EEM-ART consists of multilayer ART networks to extract novel events and encode spatio-temporal connection as episodes by incrementally generating cognitive neurons. The model connects episodes to construct a sensorimotor map for the robot to continuously perform path planning and goal navigation. Experimental results for a mobile robot indicate that EEM-ART can process multiple sensory sources for learning events and encoding episodes simultaneously. The model overcomes perceptual aliasing and robot localization by recalling the encoded episodes with a new anticipation function and generates sensorimotor map to connect episodes together to execute tasks continuously with little to no human intervention.

An Initiative Service Method Based on Fuzzy Analytical Hierarchy Process and Context Intention Inference for Drinking Service Robot

Man Hao, Wei-Hua Cao, Min Wu, Zhen-Tao Liu, Si-Han Li

To ensure that people can replenish the right amount of water when they are short of it, an initiative service method based on the fuzzy analytical hierarchy process (FAHP) and context intention inference for the drinking service robot is proposed, in which the factors that affect people's drinking demand including individual factors (e.g., physical activity, age, and gender) and environmental factors (e.g., temperature and humidity) are taken into account for drinking service. In addition, a two-level hierarchy structure for human intention together with intention degree is proposed, in which weights between context features and intention are calculated by FAHP, which imitates human's subjective judgment process mathematically and mentally. Moreover, the relationship between human intention and their demands is established by a demand analysis model for providing satisfactory service of robots. Drinking service experiments are performed in a laboratory scenario using a humans-robots interaction system, from which the experimental results show that the resultant p-value in Condition 1 is 0.455, and the resultant p-value in Condition 2 is 0.09, and the averaged values of satisfaction feedback are 2.15 and 1.85 (satisfaction from 1 to 5 decreased in turn), respectively, approximately satisfied, which demonstrates the feasibility of the proposal.

Fuzzy Linguistic Odor Cognition for Robotics Olfaction

Dapeng Yan, Hui Cao, Panpan Zhang, Shuo Yang

The robotic olfaction helps the robot to measure the odor and provides it the environment recognition

ability. This paper proposes a fuzzy linguistic odor cognition method for robotics olfaction. The language is the presentation mode and the information medium of human mind. The core idea of the proposed method is that the fuzzy linguistic rules are used to realize the brain-inspired knowledge representation. The proposed method is a cognitive model which consists of a set of fuzzy linguistic rules. The input variables of the rule antecedents are the signals of different gas sensors on the robotic electronic nose. The rule consequent is the odor label. The fuzzy linguistic rules are initialized by the density-based spatial clustering of applications with noise. The parameters of the membership functions and the rule consequents are estimated by an iterative optimization process. The experiments are performed on two real robotics olfaction data sets obtained under uncontrolled realistic conditions. The proposed method is compared with decision tree, k-nearest neighbor (KNN), support vector machine (SVM), and backpropagation neural network (BPNN). The experimental results verify that the effectiveness of the proposed method is better than that of decision tree, KNN, SVM, and BPNN.

Evolving a Sensory–Motor Interconnection Structure for Adaptive Biped Robot Locomotion

Azhar Aulia Saputra, János Botzheim, Naoyuki Kubota

We present an evolving neural oscillator-based bio-inspired biped robot locomotion for minimizing the constraints during the locomotion process. Sensory-motor coordination model is represented by the interconnection between motor neurons and sensory neurons. An evolutionary computation technique is applied for reconstructing the number of joints and the number of neurons in each joint depending on the environmental condition. In this system, either the number of joints, or the number of neurons, or the interconnection structure are dynamically changed depending on the conditions acquired from the sensors that equipped in the robot. Bacterial programming is inspired by the evolutionary process of bacteria, including bacterial mutation and gene transfer. This system is applied in computer simulation for realizing the optimization process and the optimized structure is applied in a small humanoid robot. In experiments, we run the robot in several different environmental conditions. Different neuron structures are resulted depending on the environmental conditions. The proposed tree structurebased optimization strategy can simplify the sensory-motor interconnection structure.

Regrasp Planning Using Stable Object Poses Supported by Complex Structures

Jiayao Ma, Weiwei Wan, Kensuke Harada, Qiuguo Zhu, Hong Liu

Using regrasp planning, a robot could pick up an object, place it down to an intermediate stable state, and reorient the object into certain poses by grasping and picking it up again. Regrasp is an important skill when a robot cannot reorient the object directly with one grasp due to kinematic constraints and collisions. It uses intermediate object states to release, regrasp, and reorient objects. In our previous work, we developed regrasp algorithms considering intermediate stable states on simple fixtures like a flat table surface. This paper further develops our previous studies by enabling regrasp planning using stable object poses on complex structures. The complex structures have high variety of contact elements. They not only provide flat surface supports but also point supports, line supports, and a combination of them. In detail, the developed regrasp planner includes two parallel processes. One is a dynamic simulator that computes immediate stable poses on given supporting structures. The other builds a regrasp graph using the stable states and finds a sequence of reorient motion by searching the graph. We performed thousands of simulation and real-world verification to analyze the performance of the developed planner. We conclude that our planner has higher performance on regrasping than the previous ones. The number of candidate regrasp sequences increased and the lengths became shorter. We also compared simulation results with real-world executions, and gave suggestions on the selection of supporting structures considering real-world implementations.

Using High-Frequency Local Field Potentials From Multicortex to Decode Reaching and Grasping Movements in Monkey

Peng Zhang, Jian Huang, Wei Li, Xuan Ma, Peipei Yang, Jun Dai, Jiping He

Intracortical brain-machine interfaces (iBMIs) hold the promise to restore communication and movement ability of paralyzed people. Recent studies showed that local field potentials (LFPs) could be a reliable neural signal for movement intention decoding in iBMIs. However, previous studies investigated primarily on low-frequency LFPs, while the LFPs used were recorded from only one or two cortices. The aim of this paper is to investigate the potential of high-frequency LFPs (200-300 Hz) from multicortex in movement intention decoding. In this paper, LFPs were recorded via microelectrode arrays chronically implanted into the primary motor cortex (M1), somatosensory cortex (S1), and posterior parietal cortex of two monkeys while they were trained to perform 3-D reaching and grasping movements. The wavelet packet transform (WPT) method was used to extract the time and frequency information of the high-frequency LFP signals and the node energy of the WPT coefficients was selected as the features. After feature reduction by the principal component analysis, a support vector machine decoder was used to classify discrete reaching positions and grasping postures. Our results indicate that high decoding accuracy can be achieved by the high-frequency LFPs with WPT method and this kind of LFPs could serve as useful signals in iBMIs for movement intention decoding. Moreover, better and more robust decoding performance can be achieved

by LFPs from multicortex than single-cortex.

Biologically Inspired Motion Modeling and Neural Control for Robot Learning From Demonstrations

Chenguang Yang, Chuize Chen, Ning Wang, Zhaojie Ju, Jian Fu, Min Wang

In this paper, we propose a biologically inspired framework for robot learning based on demonstrations. The dynamic movement primitive (DMP), which is motivated by neurobiology and human behavior, is employed to model a robotic motion that is generalizable. However, the DMP method can only be used to handle a single demonstration. To enable the robot to learn from multiple demonstrations, the DMP is combined with the Gaussian mixture model (GMM) to integrate the features of multiple demonstrations, where the conventional GMM is further replaced by the fuzzy GMM (FGMM) to improve the fitting performance. Also, a novel regression algorithm for FGMM is derived to retrieve the nonlinear term of the DMP. Additionally, a neural network-based controller is developed for the robot to track the generated motions. In this network, the cerebellar model articulation controller is employed to compensate for the unknown robot dynamics. The experiments have been performed on a Baxter robot to demonstrate the effectiveness of the proposed methods.

A Reinforcement Learning Architecture That Transfers Knowledge Between Skills When Solving Multiple Tasks

Paolo Tommasino, Daniele Caligiore, Marco Mirolli, Gianluca Baldassarre

When humans learn several skills to solve multiple tasks, they exhibit an extraordinary capacity to transfer knowledge between them. We present here the last enhanced version of a bio-inspired reinforcement-learning (RL) modular architecture able to perform skill-to-skill knowledge transfer and called transfer expert RL (TERL) model. TERL architecture is based on a RL actor-critic model where both actor and critic have a hierarchical structure, inspired by the mixture-of-experts model, formed by a gating network that selects experts specializing in learning the policies or value functions of different tasks. A key feature of TERL is the capacity of its gating networks to accumulate, in parallel, evidence on the capacity of experts to solve the new tasks so as to increase the responsibility for action of the best ones. A second key feature is the use of two different responsibility signals for the experts' functioning and learning: this allows the training of multiple experts for each task so that some of them can be later recruited to solve new tasks and avoid catastrophic interference. The utility of TERL mechanisms is shown with tests involving two simulated dynamic robot arms engaged in solving reaching tasks, in particular a planar 2-DoF arm, and a 3-D 4-DoF arm.

Volume 11, Issue 3, September 2019

Two Demonstrators Are Better Than One—A Social Robot That Learns to Imitate People With Different Interaction Styles

Phoebe Liu, Dylan F. Glas, Takayuki Kanda, Hiroshi Ishiguro

With recent advances in social robotics, many studies have investigated techniques for learning top-level multimodal interaction logic by imitation from a corpus of human-human interaction examples. Most such studies have taken the approach of learning equally from a variety of demonstrators, with the effect of reproducing a mixture of their average behavior. However, in many scenarios it would be desirable to reproduce specific interaction styles captured from individuals. In this paper, we train one deep neural network jointly on two separate corpuses collected from demonstrators with differing interaction styles. We show that training on both corpuses together improves performance in terms of generating socially appropriate behavior even when reproducing only one of the two styles. Furthermore, the trained neural network also enables us to synthesize new interaction styles on a continuum between the two demonstrated interaction styles. We discuss plots of the hidden layer activations from the neural network, indicating the types of semantic information that appear to be learned by the system. Further, we observe that the better performance with the synthesized corpus is not merely due to the increase of the sample size, as even with the same number of training examples, training on half the data from each corpus provided better performance than training on all the data from a single corpus.

A Personalized and Platform-Independent Behavior Control System for Social Robots in Therapy: Development and Applications

Hoang-Long Cao, Greet Van de Perre, James Kennedy, Emmanuel Senft, Pablo Gómez Esteban, Albert De Beir, Ramona Simut, Tony Belpaeme, Dirk Lefeber, Bram Vanderborght

Social robots have been proven beneficial in different types of healthcare interventions. An ongoing trend is to develop (semi-)autonomous socially assistive robotic systems in healthcare context to improve the

level of autonomy and reduce human workload. This paper presents a behavior control system for social robots in therapies with a focus on personalization and platform-independence. This system architecture provides the robot an ability to behave as a personable character, which behaviors are adapted to user profiles and responses during the human-robot interaction. Robot behaviors are designed at abstract levels and can be transferred to different social robot platforms. We adopt the component-based software engineering approach to implement our proposed architecture to allow for the replaceability and reusability of the developed components. We introduce three different experimental scenarios to validate the usability of our system. Results show that the system is potentially applicable to different therapies and social robots. With the component-based approach, the system can serve as a basic framework for researchers to customize and expand the system for their targeted healthcare applications.

BioVision: A Biomimetics Platform for Intrinsically Motivated Visual Saliency Learning

Céline Craye, David Filliat, Jean-François Goudou

We present BioVision, a bio-mimetics platform based on the human visual system. BioVision relies on the foveal vision principle based on a set of cameras with wide and narrow fields of view. We present in this platform a mechanism for learning visual saliency in an intrinsically motivated fashion. This model of saliency, learned and improved on-the-fly during the robot's exploration provides an efficient tool for localizing relevant objects within their environment. The proposed approach includes two intertwined components. On the one hand, a method for learning and incrementally updating a model of visual saliency from foveal observations. On the other hand, we investigate an autonomous exploration technique to efficiently learn such a saliency model. The proposed exploration, based on the intelligent adaptive curiosity (IAC) algorithm is able to drive the robot's exploration so that samples selected by the robot are likely to improve the current model of saliency. We then demonstrate that such a saliency model learned directly on a robot outperforms several state-of-the-art saliency techniques, and that IAC can drastically decrease the required time for learning a reliable saliency model. We also investigate the behavior of IAC in a non static environment, and how well this algorithm can adapt to changes.

Personalized Robot Assistant for Support in Dressing

Céline Craye, David Filliat, Jean-François Goudou

Robot-assisted dressing is performed in close physical interaction with users who may have a wide range of physical characteristics and abilities. Design of user adaptive and personalized robots in this context is still indicating limited, or no consideration, of specific user-related issues. This paper describes the development of a multimodal robotic system for a specific dressing scenario-putting on a shoe, where users' personalized inputs contribute to a much improved task success rate. We have developed: 1) user tracking, gesture recognition, and posture recognition algorithms relying on images provided by a depth camera; 2) a shoe recognition algorithm from RGB and depth images; and 3) speech recognition and text-to-speech algorithms implemented to allow verbal interaction between the robot and user. The interaction is further enhanced by calibrated recognition of the users' pointing gestures and adjusted robot's shoe delivery position. A series of shoe fitting experiments have been performed on two groups of users, with and without previous robot personalization, to assess how it affects the interaction performance. Our results show that the shoe fitting task with the personalized robot is completed in shorter time, with a smaller number of user commands, and reduced workload.

A Wireless Multifunctional SSVEP-Based Brain–Computer Interface Assistive System

Chin-Teng Lin, Ching-Yu Chiu, Avinash Kumar Singh, Jung-Tai King, Li-Wei Ko, Yun-Chen Lu, Yu-Kai Wang

Several kinds of brain–computer interface (BCI) systems have been proposed to compensate for the lack of medical technology for assisting patients who lose the ability to use motor functions to communicate with the outside world. However, most of the proposed systems are limited by their nonportability, impracticality, and inconvenience because of the adoption of wired or invasive electroencephalography acquisition devices. Another common limitation is the shortage of functions provided because of the difficulty of integrating multiple functions into one BCI system. In this paper, we propose a wireless, noninvasive and multifunctional assistive system which integrates steady state visually evoked potential-based BCI and a robotic arm to assist patients to feed themselves. Patients are able to control the robotic arm via the BCI to serve themselves food. Three other functions: 1) video entertainment; 2) video calling; and 3) active interaction are also integrated. This is achieved by designing a functional menu and integrating multiple subsystems. A refinement decision-making mechanism is incorporated to ensure the accuracy and applicability of the system. Fifteen participants were recruited to validate the usability and performance of the system. The averaged accuracy and information transfer rate achieved is 90.91% and 24.94 bit per min, respectively. The feedback from the participants demonstrates that this assistive system is able to significantly improve the quality of daily life.

Deep Spiking Convolutional Neural Network Trained With Unsupervised Spike-Timing-Dependent Plasticity

Chankyu Lee, Gopalakrishnan Srinivasan, Priyadarshini Panda, Kaushik Roy

Spiking neural networks (SNNs) have emerged as a promising brain inspired neuromorphic-computing paradigm for cognitive system design due to their inherent event-driven processing capability. The fully connected (FC) shallow SNNs typically used for pattern recognition require large number of trainable parameters to achieve competitive classification accuracy. In this paper, we propose a deep spiking convolutional neural network (SpiCNN) composed of a hierarchy of stacked convolutional layers followed by a spatial-pooling layer and a final FC layer. The network is populated with biologically plausible leaky-integrate-and-fire (LIF) neurons interconnected by shared synaptic weight kernels. We train convolutional kernels layer-by-layer in an unsupervised manner using spike-timingdependent plasticity (STDP) that enables them to self-learn characteristic features making up the input patterns. In order to further improve the feature learning efficiency, we propose using smaller 3×3 kernels trained using STDP-based synaptic weight updates performed over a mini-batch of input patterns. Our deep SpiCNN, consisting of two convolutional layers trained using the unsupervised convolutional STDP learning methodology, achieved classification accuracies of 91.1% and 97.6%, respectively, for inferring handwritten digits from the MNIST data set and a subset of natural images from the Caltech data set.

Automatic Object Searching and Behavior Learning for Mobile Robots in Unstructured Environment by Deep Belief Networks

Jiru Wang, Vui Ann Shim, Rui Yan, Huajin Tang, Fuchun Sun

Automatic object searching is one of the essential skills for domestic robots to operate in unstructured human environments. It involves concatenation of several capabilities, including object identification, obstacle avoidance, path planning, and navigation. In this paper, we propose an automatic object searching framework for a mobile robot equipped with a single RGB-D camera. The obstacle avoidance is achieved by a behavior learning algorithm based on deep belief networks. The target object is recognized using scale-invariant feature transform descriptors and the relative position between the target and mobile robot is estimated from the RGB-D data. Subsequently, the mobile robot makes a path planning to the target location using an improved bug-based algorithm. The framework is tested in indoor environments and requires the robot to perform obstacle avoidance and automatically search and approach the target object. The results indicate that the system is collision free and reliable in performing searching tasks. This system's functions make itself have the potential of being used for local navigation in unstructured environments.

Adaptive Behavior Acquisition of a Robot Based on Affective Feedback and Improvised Teleoperation

Masakazu Hirokawa, Atsushi Funahashi, Yasushi Itoh, Kenji Suzuki

In socially assistive robotics, especially for children with autism spectrum disorder (ASD), adapting the behavior of the robot according to the personal characteristics of each individual is one of the important challenges. Machine learning techniques are promising approaches to endow a robot with the capability of adapting its behavior through the interaction. It is critical to prepare a rich data set such as a set of behaviors with teaching signals for each individual with ASD to allow application of the state-of-the-art machine learning techniques; however, this is typically difficult to prepare in advance owing to the diverseness of ASD and the complexity of the motion design of the robot. This paper proposes a framework to acquire the personalized behavior set of a robot by combining a robot teleoperation method and a wearable device for detecting the affective cue of a child with ASD while interacting with the robot. The developed system allows the human operator to improvise the robot's behavior flexibly in real-time to explore the preferred interaction manner and motion patterns of each child. The preferred motion patterns are extracted and evaluated based on the affective state of the child estimated by the wearable device, and stored in the personal database for each individual with ASD. We conducted a free-interaction experiment with ten participants with ASD and demonstrated that the proposed system successfully described the interaction between the robot and the participant for acquiring the appropriate behaviors of the robot.

Motor-Imagery-Based Teleoperation of a Dual-Arm Robot Performing Manipulation Tasks

Yiliang Liu, Wenbin Su, Zhijun Li, Guangming Shi, Xiaoli Chu, Yu Kang, Weiwei Shang

This paper proposes a brain-computer interface (BCI)-based teleoperation strategy for a dual-arm robot carrying a common object by multifingered hands. The BCI is based on motor imagery of the human brain, which utilizes common spatial pattern method to analyze the filtered electroencephalograph signals. Human intentions can be recognized and classified into the corresponding reference commands in task space for the robot according to phenomena of event-related synchronization/desynchronization, such that the object manipulation tasks guided by human user's mind can be achieved. Subsequently, a concise dynamics consisting of the dynamics of the robotic arms and the geometrical constraints between the end-effectors and the object is formulated for the coordinated dual arm. To achieve optimization motion in the task space, a redundancy resolution at velocity level has been implemented through neural-dynamics optimization. Extensive experiments have been made by a number of subjects, and the results were provided to demonstrate the effectiveness of the proposed control strategy.

BUM: Bayesian User Model for Distributed Learning of User Characteristics From Heterogeneous Information

Gonçalo S. Martins, Luís Santos, Jorge Dias

This paper presents a Bayesian user model able to learn and estimate user characteristics in a distributed manner using heterogeneous information. A unified user representation is obtained from an inference process, receiving a set of independently estimated user characteristics from different sources. The independence of characteristic models enables the system to be modular, with each module estimating one characteristic. The proposed model is iterative, fusing new observations, and measurements with previous information in a process regulated entropy. The system allows diverse implementations, such as the combination of multiple robots with a cloud infrastructure or distributed ambient sensors. This paper aims to enable the system to perform online learning while interacting with users. The system is also able to obtain a correct user representation from heterogeneous information, even when some user characteristics cannot be computed. To demonstrate its functionality, the system is tested on two experimental datasets, obtained from simulated experiments and with real users. This technique advances the state of the art in the areas of AAL and user-adaptive systems, and in cloud-connected robots and Internet of Things, allowing for these heterogeneous and naturally distributed teams of devices to better model their users, potentially achieving higher interaction autonomy.

Modeling the Co-Emergence of Linguistic Constructions and Action Concepts: The Case of Action Verbs

Maximilian Panzner, Judith Gaspers, Philipp Cimiano

In this paper, we are concerned with understanding how linguistic and conceptual structures co-emerge, shaping and influencing each other. Most theories and models of language acquisition so far have adopted a "mapping" paradigm according to which novel words or constructions are "mapped" onto existing, priorly acquired or innate concepts. Departing from this mapping approach, we present a computational model of the coemergence of linguistic and conceptual structures. We focus in particular on the case of action verbs and develop a model by which a system can learn the grounded meaning of a verbal construction without assuming the prior existence of a corresponding sensomotorically grounded action concept. Our model spells out how a learner can distill the essence of the meaning of a verbal construction as a process of incremental generalization of the meaning of action verbs, starting from a meaning that is specific to a certain situation in which the verb has been encountered. We understand the meaning of verbs as evoking a grounded simulation rather than a static concept and propose to capture the meaning of verbs via generative statistical models that support simulation, in our case hidden Markov models. Statistical models can represent the essence of a verb's meaning while modeling uncertainty and thus variation at the surface level of (observed) action performances. We show that by extending an existing framework for construction learning, our approach can account for the co-emergence of linguistic and conceptual structures. We provide proof-of-concept for our model by experimentally evaluating it on matching, choice, and generation tasks, showing that our model can not only understand but also produce language.

A System for Noncontact Estimation of Cognitive Load Using Saccadic Parameters Based on a Serio-Parallel Computing Framework

Anirban Dasgupta, Som Madhab Bhattacharya, Aurobinda Routray

This paper proposes a system which uses a three-stage serio-parallel video-oculographic framework for computing the saccadic eye parameters to indicate the amount of cognitive loading. The three stages are viz., face and eye detection, iris and eye corner localization, and finally saccadic parameter computation. Since saccades are fast movements of the eyeballs, accurate estimation of these parameters requires high frame rates of acquisition and processing. Our proposed framework meets such deadlines by accelerating the process using graphics processing units (GPUs). The first stage comprises the face and eye detection using the respective Haar classifiers followed by tracking of a region of interest using a minimum output sum of squared error filter. In the second stage, the filter parameters are transferred to the GPU, where our proposed parallel scheme is implemented. In the detected eye region, the iris candidates are ranked using a sum of dot products of normalized displacement vectors with gradient vectors. We also localize the eye corners as the reference points. The saccadic velocity and duration are obtained using this eye position signal in the third stage. Finally, the amount of cognitive loading is determined based on these parameters.

EDITOR **Pierre-Yves Oudeyer,** Inria and Ensta ParisTech, France, pierre-yves.oudeyer@inria.fr
EDITORIAL ASSISTANT **Fabien C. Y. Benureau,** CNRU, Okinawa Institute of Science and Technology (OIST), Japan
ISSN 1550-1914