Berthouze, L., Kaplan, F., Kozima, H., Yano, H., Konczak, J., Metta, G., Nadel, J., Sandini, G., Stojanov, G. and Balkenius, C. (Eds.)
Proceedings of the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems
Lund University Cognitive Studies, 123. ISBN 91-974741-4-2

Using social robots to study abnormal social development

Brian Scassellati
Department of Computer Science
51 Prospect Street
New Haven, CT 06520
scaz@cs.yale.edu

Abstract

Social robots recognize and respond to human social cues with appropriate behaviors. Social robots, and the technology used in their construction, can be unique tools in the study of abnormal social development. Autism is a pervasive developmental disorder that is characterized by social and communicative impairments. Based on three years of integration and immersion with a clinical research group which performs more than 130 diagnostic evaluations of children for autism per year, this paper discusses how social robots will make an impact on the ways in which we diagnose, treat, and understand autism.

1. Introduction

For the past three years, our robotics group has been immersed in one of the premiere clinical research groups studying autism, led by Ami Klin and Fred Volkmar at the Yale Child Study Center. This paper outlines the conclusions that we have reached in our attempts to apply technology from social robotics to the unique clinical problems of autism.

Section 2 provides an introduction to autism which highlights some of the difficulties with current diagnostic standards and research techniques. Section 3 describes attempts to use robots as therapeutic aids and discusses the as yet unfulfilled promise of these methods. Section 4 describes how diagnosis can be improved through the use of both passive social cue measurement and interactions with a social robot to provide quantitative, objective measurements of social response. Section 5 speculates on how the use of social robots in autism research might lead to a greater understanding of the disorder.

2. What we know about autism

Autism was first identified in 1943 by Kanner who emphasized that this congenital condition was characterized by an inability to relate to other people from the first days of life. Current research suggests that 1 in every 300 children will be diagnosed with the broadly-defined autism spectrum disorder (ASD), but studies have found prevalence rates that vary between 1 in every 500 to 1 in every 166. For comparison, 1 in every 800 children is born with Down syndrome, 1 in every 450 will have juvenile diabetes, and 1 in every 333 will develop cancer by the age of 20. Furthermore, the rate of diagnosis increased six-fold between 1994 and 2003. It is unclear how much of this increase is a result of changes in the diagnostic criteria, increases in awareness, or a true increase in prevalence [1].

The social disability in autism is a profound one affecting a person's capacity for understanding other people and their feelings, and for establishing reciprocal relationships. To date, autism remains a behaviorally specified disorder [2]; there is no blood test, no genetic screening, and no functional imaging test that can diagnose autism. Diagnosis relies on the clinician's intuitive feel for the child's social skills including eye-to-eye gaze, facial expression, body postures, and gestures. These observational judgments are then quantified according to standardized protocols, e.g. [3,4] that are both imprecise and subjective. The broad disagreement of clinicians on individual diagnoses creates difficulties both for selecting appropriate treatment for individuals and for reporting the results of population-based studies [5,6]. Genetic studies have underscored the importance of understanding both the broader phenotype of autism and the remarkable heterogeneity in syndrome expression. However, the causes and etiology of the disorder are still unknown [2]. A more precise characterization and quantification of social dysfunction is required to direct neurobiological research in autism is still lacking [7,8,9].

3. Sources of motivation and incremental adaptation for therapy

A few projects world-wide seek to include robots as part of the therapeutic regimen for individuals with autism [10,11,12,13]. The motivation for much of this work comes from the observation that robots generate a high degree of motivation and engagement in subjects, including subjects who are unlikely or unwilling to interact socially with human therapists. Initial results show that robots can serve as a medium of communication between individuals with autism and their caregivers but none have yet shown transfer of skills from a robot partner to a human partner.

Within our own lab, these results have been confirmed with an extremely simple commercial robot called ESRA which generates a small set of facial expressions using five servos. In one experiment, ESRA was programmed to perform a short script which included both a set of actions and an accompanying audio file that was played from speakers hidden near the robot. 13 subjects (mean age 3.4 years) including 7 children with autism spectrum disorders and 6 typically developing children were positioned across a table from ESRA. The script started with the robot "waking up", asking a few questions of the child, and then falling back "asleep". The robot had no sensory capabilities and did not respond to anything that the child did. Even with the extremely limited capabilities of ESRA, the robot was well tolerated by all of the children and many of them (including many of those within the autism spectrum) seemed to thoroughly enjoy the session. Children were universally engaged with the robot, and often spent the majority of the session touching the robot, vocalizing at the robot, and smiling at the robot. It is worth noting that for many of the children with autism in this pilot study, these positive proto-social behaviors are rarely seen in a naturalistic context.

The great hope of this line of research is the development of a "social crutch," a robot that motivates and engages children, teaches them social skills incrementally, and assists in the transfer of this knowledge to interactions with humans. The robot's behavior can be decomposed arbitrarily, turning off some behaviors while leaving others intact, allowing the selective construction complex social abilities through layers of social responses, sometimes in combinations that cannot be performed by humans. This layering of response allows the therapist to focus on single behaviors while ignoring all other social factors or maintaining their response at a constant. This type of isolation of cues and responses is difficult to train human therapists to perform. The as yet unfulfilled promise of this line of research is that learning skills with a robot will be simpler because of the ability to isolate particular responses, thus allowing a unique form of incremental therapy. In a different domain, but using a similar principle, data suggests that computerized face perception training leads to therapeutic benefits for individuals with autism [14].

4. Quantitative, objective metrics for diagnosis

Many of the diagnostic problems associated with autism would be alleviated by the introduction of quantitative, objective measurements of social response. We believe that this can be accomplished through two methods: through passive observation of the child at play or in interactions with caregivers and clinicians, and through structured interactions with robots that are able to create standardized social "presses" designed to elicit particular social responses. While the information gathered from both passive and interactive systems will not replace the expert judgment of a trained clinician, providing high-reliability quantitative measurements will provide a unique window into the way in which children with autism attempt to process naturalistic social situations. These metrics provide both an opportunity to compare populations of individuals in a standardized manner and the possibility of tracking the progress of a single individual across time. Because some of the social cues that we measure (gaze direction in particular) are recorded in greater detail and at an earlier age than can occur in typical clinical evaluations, one possible outcome of this work is a performance-based screening technique capable of detecting vulnerability for autism in infants and toddlers.

4.1 Passive Sensing

Passive sensors record information on social response without directly engaging in interactions. In many cases, the perceptual systems of a social robot can act as a passive social cue sensor. To evaluate the usefulness of this idea, we have outfitted some of our clinical evaluation rooms with cameras and microphones and software similar to that used on the social robots Nico, Cog, and Kismet [15,16,17]. Most of these passive sensors record and interpret data while the subjects are actively engaged in standard clinical evaluations and do not require any specific protocol to be employed. Three examples follow.

Gaze direction and focus of attention: For several years, we have used commercial eye-tracking systems which require subjects to wear a baseball cap with an inertial tracking system and camera/eyepiece assembly which allows us to record close-up images of one eye. When viewing naturalistic social scenes, adolescents and adults with autism display gaze patterns which differ significantly between control populations [9,18,19]. Fixation time variables predicted level of social competence (e.g., at an average r=.63). This was the first experimental measure to successfully predict level of social competence in real life for individuals with autism. Although visual fixation on regions of interest are sensitive measures of social dysfunction, moment-by-moment scan-paths are even more sensitive and offer further insight into the underlying dysfunction (see section 5 for an example) [9].

Position tracking: Some of the most basic information on social response can be derived from the relative positioning of individuals. How close a child stands in relation to an adult, how often the child approaches an adult, how much time is spent near an adult, and whether or not the child responds when an adult approaches are a few of the relatively simple statistics that can be derived from positional information. These social cues, especially the concept of "personal space," are often deficient in individuals with autism and are part of the diagnostic criteria [2]. Using a pair of calibrated stereo cameras and a computational vision system developed in part by our team [20], we have been able to successfully track the position of individuals as they move about in our clinical space. We are currently evaluating this data as predictors of social ability and eventual diagnosis.

Vocal prosody: Individuals with autism often have difficulty both generating and recognizing vocal prosody and intonation [21]. (Simply put, prosody refers to not what is said, but how it is said.) There are no standardized measures of prosody in the clinical literature [22], and the only research instrument available [23] is very laborious and thus seldom used in diagnostic evaluation or experimental studies. We recently constructed a multi-stage Bayesian classifier capable of distinguishing between five categories of prosodic speech (prohibition, approval, soothing, attentional bids, and neutral utterances) with an accuracy of more than 75% on a difficult set of vocal samples taken from typical adults (both male and female). In comparison, human judges were able to correctly classify utterances 90% of the time within this data set [24]. This system is currently being applied within our clinic.

4.2 Interactive Social Cue Measurement

While there is a vast array of information that can be obtained by passive sensing technologies, the use of interactive robots provides unique opportunities for examining social responses in a level of detail that has not previously been available. There are three key advantages to this approach. First, by generating a social press designed to elicit a particular social response from the subject, the interactive system can selectively probe for information on low-occurrence social behaviors or on behaviors that may not easily emerge in diagnostic sessions in the clinic. Second, the robot provides a repeatable, standardized stimulus and recording methodology. Because both the production and recognition are free from subjective bias, the process of comparing data on social responses between individuals or for a single individual across time will be greatly simplified. As a result, the interactive system may prove to be a useful evaluation tool in measuring the success of therapeutic programs and may provide a standard for reporting social abilities within the autism literature. Third, because a robotic system can generate social cues and record measurements autonomously, simple interactive toys can be designed to collect data outside of the clinic, effectively increasing both the quantity and quality of data that a clinician can obtain without extensive field work.

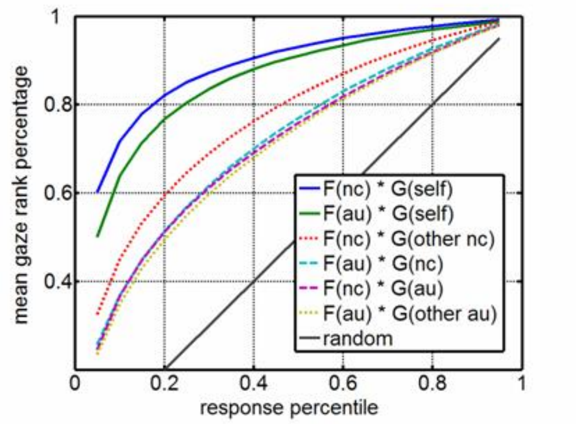

Figure 1: Results of linear discriminant analysis of autistic (au) and normal (nc) gaze patterns. Linear filters F(x) are trained to reproduce the gaze pattern G(x) of each individual x. Filters can then be applied to predict the gaze patterns of any other individual. For example, F(A)*G(self) indicates a filter trained on an individual with autism is tested on that same individual while F(NC)*G(A) indicates a filter trained on a control individual is tested on data from an individual with autism.

We have developed one simple device, called Playtest, for determining auditory preferences in the clinic or in the home. When a button is pressed, the device plays one of two audio clips, produces a series of flashing lights to entice attention, and records the time, date, button pressed and audio clip played to non-volatile memory. This device can be sent home with a family to collect information on typical play patterns. This method has been shown to have important diagnostic value since it can measure listening preferences to speech sounds, abnormalities of which are among the most robust predictors of subsequent diagnosis of autism [25].

5. Robots as tools of understanding

The fine-grained analyses of social capabilities that result from work on therapeutic and diagnostic applications have the potential to enhance our understanding of autistic disorders. We have already encountered one example of this potential in our pilot studies of gaze detection. Based on our earlier observations on the differences in gaze direction between typically developing individuals and individuals with autism and in response to our need to characterize potential looking patterns for a robot, we have begun to generate predictive models that show not only the focus of an individual's gaze but also provides an explanation of why they choose to look at particular locations. A simple classifier (a linear discriminant) was trained to replicate the gaze patterns of a particular individual. The performance of this predictor for a single frame is evaluated by having the filter rank-order each location in the image and selecting the rank of the location actually chosen by a particular individual. Thus, random performance across a sequence of images results in a median rank score of 50th percentile, while perfect performance would result in a median rank score of 1.0 (100th percentile). Trained filters predict the gaze location of the individual they were trained upon with very high accuracy (median rank scores of 90th -92nd percentile). By applying a filter trained on one individual to predict the data of a second individual, we can evaluate the similarity of the underlying visual search methods used by each individual. In a pilot experiment with this technique, typically developing individuals were found to all use similar strategies (median rank score in the 86th percentile). Significantly, autistic individuals failed to show similar visual search strategies both among other individuals with autism (73rd percentile) and among the typically developing population (72nd percentile). Filters trained on our control population were similarly unsuccessful at predicting the gaze patterns of individuals with autism (71st percentile). Significant differences (all p¡0.01 for a two-tailed t-test) are seen between the following classes: (1) F(NC)*G(self), (2) F(A)*G(self), (3) F(NC)*G(NC), and (4) the three other conditions. These preliminary results suggest that while our control population all used some of the same visual search strategies, individuals with autism were both not consistently using the same strategies as the control population nor were they using the strategies that other individuals with autism used.

Acknowledgements

This work would not be possible without the patience and innovation of Ami Klin, Fred Volkmar, and their entire team at the Child Study Center. Ganghua Sun, Fred Shic, Liz Darbie, Avi Robinson-Mosher, Jim Logan, and Reuben Grinberg contributed to some of the preliminary findings reported here.

References:

[1] Centers for Disease Control and Prevention, http://www.cdc.gov/ncbddd/dd/ddautism.htm, May 3, 2005.

[2] Volkmar, F.R., Lord, C., Bailey, A., Schultz, R.T., & Klin, A. (2004). Autism and pervasive developmental disorders. Journal of Child Psychology and Psychiatry, 45(1), 1-36.

[3] Sparrow, S.S., Balla, D., & Cicchetti, D. (1984). Vineland Adaptive Behavior Scales, Expanded Edition. Circle Pines, MN: American Guidance Service.

[4] Mullen, E.M. (1995) Mullen Scales of Early Learning: AGS Edition. Circle Pines, MN: AGS.Mullen, 1995.

[5] Klin, A., Lang, J., Cicchetti, D.V., & Volkmar, F.R. (2000). "Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial." Journal of Autism and Developmental Disorders, 30(2), 163-167.

[6] Volkmar, F.R., Chawarska, K., & Klin, A. (2005). "Autism in infancy and early childhood." In A. Kazdin (Ed.), Annual Review of Psychology, 56, 315-36.

[7] Schultz, R.T., & Robins, D. (2005). "Functional neuroimaging studies of autism." In F.R. Volkmar, R. Paul, A. Klin, & D.J. Cohen (Eds.), Handbook of Autism and Pervasive Developmental Disorders, 3rd edition. New York: Wiley.

[8] Bailey, A., Phillips, W., and Rutter, W. (1996). "Autism: Towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives." Journal of Child Psychology and Psychiatry, 37: 89-126.

[9] Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002a). "Defining and quantifying the social phenotype in autism." American Journal of Psychiatry, 159(6), 895-908.

[10] Werry, I.P., Dautenhahn, K. (1999). "Applying Mobile Robot Technology to the Rehabilitation of Autistic Children." Proceedings of SIRS '99, Symposium on Intelligent Robotics Systems, 20-23 July 1999.

[11] Michaud, F., Thberge-Turmel, C. (2002), "Mobile robotic toys and autism", Socially Intelligent Agents - Creating Relationships with Computers and Robots, Kerstin Dautenhahn, Alan Bond, Lola Canamero, Bruce Edmonds (editors), Kluwer Academic Publishers, pages 125-132.

[12] Dautenhahn, K. (2000) "Design Issues on Interactive Environments for Children with Autism." Proceedings International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT), p. 153-161.

[13] Kozima, H., Nakagawa, C., and Yano, H. "Designing a Robot for Spatio-Temporal Contingency-Detection Game," International Workshop on Robotic and Virtual Agents in Autism Therapy (Hospital la Salpetriere, Paris, France), 2002.

[14] Schultz, R. Personal Communication.

[15] Scassellati, B. (2003) "Investigating models of social development using a humanoid robot" International Joint Conference on Neural Networks, CD-ROM.

[16] Brooks, R. Breazeal, C., Marjanovic, M., Scassellati, B., Williamson, M. (1998) "The Cog Project: Building a Humanoid Robot." In C. Nehaniv, ed., Computation for Metaphors, Analogy and Agents, Vol. 1562 of the Springer Lecture Notes in Artificial Intelligence. P. 52-87.

[17] Breazeal, C., Edsinger, A., Fitzpatrick, P., Scassellati, B. (2001) "Active vision for sociable robots." IEEE Transactions on Systems, Man, and Cybernetics,Vol. 31, No. 5, p. 443-453.

[18] Klin, A., Jones, W., Schultz, R., Volkmar, F.R., Cohen, D.J. (2002b). "Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism," Archives of General Psychiatry, 59(9), 809-816.

[19] Klin, A., Jones, W., Schultz, R.T., & Volkmar, F.R. (2003). "The Enactive Mind - from actions to cognition: Lessons from autism." Philosophical Transactions of the Royal Society, Biological Sciences, 358, 345-360.

[20] Scassellati, B. (2002). "Theory of mind for a humanoid robot." Autonomous Robots, 12, 13-24.

[21] Shriberg, L.D., Paul, R., McSweeny, J.L., Klin, A., Cohen, D.J., & Volkmar, F.R. (2001). "Speech and prosody characteristics of adolescents and adults with high functioning autism and Asperger syndrome." Journal of Speech, Language, and Hearing Research, 44, 1097-1115.

[22] Paul, R. (2005). "Communicative competence in individuals with autism". In F.R. Volkmar, R. Paul, A. Klin, & D.J. Cohen (Eds.), Handbook of Autism and Pervasive Developmental Disorders, 3rd edition. New York: Wiley.

[23] Shriberg, L. Kwiatkowski, J., & Rasmussen, C. (1990). Prosody-Voice Screening Profile. Tuscon, AZ: Communication Skillbuilders.

[24] Robinson-Mosher, A., Scassellati, B. (2004) "Prosody recognition in male infant-directed speech." Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Sendai, Japan.

[25] Klin, A. (1991). "Young autistic Children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal." Journal of Autism and Developmental Disorders, 21(1), 29-42.