Attachment Theory in Long-Term Human-Robot Interaction

Braden McDorman, Caitlyn Clabaugh, and Maja J Matarić
Viterbi School of Engineering
University of Southern California
Los Angeles, California 90089–0781
{mcdorman, clabaugh, mataric}@usc.edu

Abstract—We propose attachment theory as an important framework for long-term, human-robot interaction. In this paper, we provide background on attachment theory from social psychology research pertaining to human-human relationships. We then discuss the implications of such research on attachment relationships between people and robots or other artificial social agents in terms of attachment styles, formation, and dynamics.

Keywords—attachment theory, human-robot interaction, socially assistive robotics, long-term.

I. Introduction

Attachment theory is the formalization of *attachment*, including its styles, formation, and dynamics, in relationships between people [1]. Originally proposed to describe the relationship dynamics between a primary caregiver and her infant, attachment theory has since been expanded to describe relationships more generally. In this paper, we discuss aspects of attachment theory as they relate to human-agent relationships, specifically the applications of mature insights from psychology to novel long-term human-robot interaction (HRI) research.

As long-term and in-home HRI becomes possible with cost-effective high-quality robotic platforms, researchers and commercial vendors alike will benefit from understanding possible social relationships that could form between robots and their users. This includes attachment relationships, which are important in particular due of their implications for human health and well-being [2] emphasized in socially assistive robotics (SAR). Thus, attachment theory has implications to personal robotics as well as other artificial agents in various novel therapeutic use-cases, such as relationship counseling or emotion-focused therapy.

In SAR and HRI more generally, attachment may be leveraged to facilitate long-term human-robot collaboration and development. Viewing long-term HRI from the vantage point of attachment theory leads to insights and conclusions that will inform the design and interaction style of future social agents. This paper gives background on attachment theory from social psychology and explores its possible implications for long-term HRI.

II. ATTACHMENT STRATEGIES AND STYLES

Attachment theory describes both a working model of the attachment behavior system, herein referred to as attachment

strategies, and an individual's overarching personal relationship dynamics, herein referred to as attachment *styles*. The repeated application of attachment strategies leads to a longer term, learned attachment style.

A. Attachment Strategies

Attachment theory categorizes the attachment behavior system in terms of affect regulation *strategies* (Fig. 1) [3]. These strategies can lead to successfully or unsuccessfully coping with anxiety-inducing triggers, such as personal conflict or an upsetting life event, and can either positively or negatively reinforce activation of the attachment system. This trigger leads to an individual seeking closeness to a relationship partner. The partner's response determines how the attachment system is affected.

- Secure Strategy The partner is attentive and responsive, which leads to effective co-regulation of anxiety and reinforcement of the activation of the attachment system.
- Attachment-Anxiety Strategy The partner is not attentive or responsive, leading to increased anxiety and hyper-activation of the attachment system, as the individual continues to seek closeness with the partner.
- Attachment-Avoidance Strategy The partner is not attentive or responsive, leading to anxiety suppression, cognitive, emotional, and physical distancing, and deactivation of the attachment system.

B. Attachment Styles

A person's attachment style is a largely static manifestation of long-term learned attachment behavior. The attachment style defines the dynamics and interaction styles that permeate most of that person's attachment relationships. While these styles exist on continuous spectra, they can be grouped into four quadrants (Fig. 2) [4], as follows.

- Secure Comfortable with intimacy and autonomy.
 Defined by a positive view of self and others.
- Preoccupied Preoccupied with relationships. Defined by a negative view of self and positive view of others.
- Dismissing Dismissing of intimacy and counterdependent. Defined by a positive view of self and negative view of others.

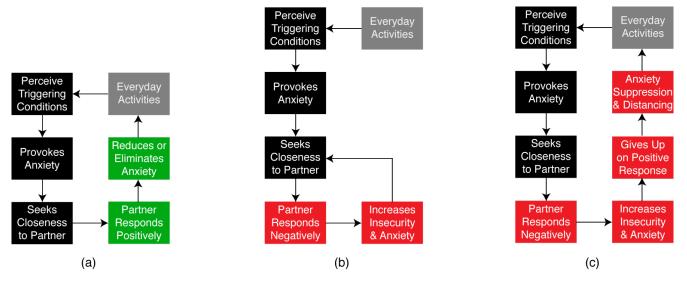


Fig. 1: Attachment strategies: (a) secure attachment, (b) attachment-anxiety, and (c) attachment-avoidance.

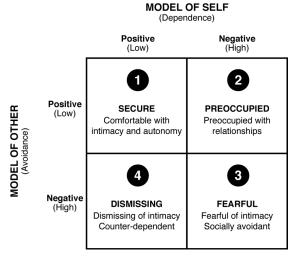


Fig. 2: Attachment styles viewed as a parametrization of an agent's model of self and other.

Fearful – Fearful of intimacy and socially avoidant.
 Defined by a negative view of self and others.

Attachment styles are learned over long periods of time and remain static for as many as 70% of individuals. The remaining 30% may change attachment styles over short time spans (weeks to months) [5].

III. ATTACHMENT FORMATION

Attachment formation describes the process involved in two individuals forming an attachment relationship over time. We review the available four-phase process model of attachment formation [6] from social psychology literature, and adapt it to inform attachment formation with an artificial social agent. The model describes attachment "states" that may be used to build a temporal model of attachment based on both conscious and subconscious interaction features.

Pre-attachment – The individuals seek proximity to one another. Conversational content is emotionally neutral, superficial, and self-enhancing. Voice quality is animated, high pitched, and emotionally aroused. Eye contact is intermittent.

Attachment-in-the-making – Individuals seek "safe haven" with one another. Conversational content is care-eliciting and contains emotional disclosure. Voice quality is soothing and may include whispering. Eye contact is frequent with protracted mutual gazing.

Clear-cut attachment – Individuals become distressed if psychologically separated for extended times. Conversational content becomes less emotional, care-eliciting, and more mundane. Voice quality becomes context specific and more typical. Eye contact becomes less frequent, but still with protracted mutual gazing.

Goal-corrected partnership – Individuals are secure with each other. Conversational content is predominately mundane. Voice quality becomes predominately typical. Eye contact becomes less frequent, deliberate, and context-specific.

While the explored and adapted process model does not include so-called backchannel communication, nonverbal features, such as subconscious nonverbal mirroring are known to be an important part of rapport formation [7], and may influence and inform attachment formation.

IV. APPLICATIONS OF ATTACHMENT THEORY IN HRI

HRI and SAR researchers and commercial robot developers can leverage the described attachment strategies and styles to inform robot behavioral design. A socially assistive robot, for example, may need to adapt its interactive therapeutic strategies behavior based on the user's attachment style and personal relationship to the robot as a social agent. A model-level understanding of attachment theory may also allow a robot to serve in novel and specialized use-cases, such as a counselor between two individuals with a preexisting attachment relationship. In this section we explore different topics of relevance to HRI and SAR.

A. Verbal and Nonverbal Behavior Features

An individual's verbal and nonverbal behavior with close relationship partners is a function of attachment style and anxiety. An anxious individual with a dismissive or fearful attachment style will employ lower levels of gazing, facial and vocal "pleasantness", and attentiveness [8]. Such individuals also use less adaptive conflict resolution strategies [9], and may have deficits in decoding their partner's nonverbal behavior [10]. Anxious individuals with secure or preoccupied attachment styles do not exhibit these features.

A robot may be able to leverage these verbal and nonverbal features during an interaction to gain an understanding of the user's attachment style or anxiety during an interaction with itself or other social partners. These signals may serve as indicators to support the diagnosis of psychological distress conditions such as anxiety, depression, and post traumatic stress disorder [11]. A robot may also imitate an attachment style during interactions to achieve a certain personality.

B. Self-disclosure

Self-disclosure is an important component of secure attachment relationships, and may be used to strengthen an attachment between two individuals if performed in a stylesensitive and time-sensitive way. While individuals with secure or preoccupied attachment styles benefit from self-disclosure, individuals with dismissive or fearful attachment styles show increases in negative affect during these situations [12]. This indicates that a robot should understand an individual's attachment style before encouraging self-disclosure in an emotionally intimate interaction. Self-disclosure that may also violate social or relationship norms actively impedes the attachment process [6].

Interactive machine learning is a computational as well as interpersonal method in human-robot and human-computer interaction (HCI) research – the agent learns directly from the human user [13]. Through interactive machine learning, a robot or agent may request or even require self-disclosure from its user. For example, interactive machine learning methods may require the elicitation of learning-sensitive information for intelligent tutoring systems [14]. More generally, HRI and HCI may focus on emotionally intimate or sensitive topics such as socially assistive robots for post-stroke rehabilitation [15] or for use in autism research [16], and virtual reality technology for rehabilitation [17].

Individuals with dismissive or fearful attachment may be averse to physical closeness with partners, including touching, holding, and caressing. Their attachment relationships are characterized by interdependence rather than emotional intimacy. These differences may be important to the proxemic control of autonomous, embodied agents that are largely tailored to the general public and not personalized to individuals' preferences [18].

C. Interaction Duration and Content

Attachment formation requires regular, long-term, and emotionally significant interactions as constrained by the previously described attachment formation process. Because emerging HRI and SAR technologies are beginning to explore

these domains, it is important to understand and leverage attachment theory in design.

As previously discussed, HRI and SAR research may indirectly deal with attachment in emotionally intimate or sensitive domains such as rehabilitation. Attachment between the robot and its user may be a means to glean psychologically-relevant information for assessment, improve adherence and collaboration between the robot and its user, and/or influence behavioral change in the user.

However, attachment formation typically occurs over long time spans of months to years. This represents a major challenge for HRI to date, as the longest HRI studies, given the complexities involved, have so far been conducted over several months [19]. One of the longest SAR studies, performed by Tapus et al., in the context of post-stroke rehabilitation therapy was six months long [20].

V. MEASURING ATTACHMENT

The attachment style between two individuals may be measured through field-standard self-reporting questionnaires such as the Experiences in Close Relationships – Relationship Structures (ECR-RS) [21]. The ECR-RS is designed to be used in a variety of interpersonal relationship types (such as parental, friendly, and romantic) and age groups, making it a strong candidate for assessment of attachment to an artificial social agent.

VI. CONCLUSION

This paper presents a case for attachment theory as an important framework in long-term HRI and SAR with clinical or rehabilitative impacts. The strategies, styles, dynamics, formation, and measures of social attachment theory may inform our emerging efforts in long-term HRI and SAR.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Expedition in Computing grant IIS-1139148 and National Science Foundation Graduate Research Fellowship DGE-1418060.

REFERENCES

- [1] J. Cassidy, Handbook of attachment: Theory, research, and clinical applications. Rough Guides, 2002.
- [2] N. L. Collins and B. C. Feeney, "An attachment theory perspective on closeness and intimacy," *Handbook of closeness and intimacy*, pp. 163–187, 2004.
- [3] R. R. Kobak, H. E. Cole, R. Ferenz-Gillies, W. S. Fleming, and W. Gamble, "Attachment and emotion regulation during mother-teen problem solving: A control theory analysis," *Child development*, pp. 231–245, 1993.
- [4] K. Bartholomew and L. M. Horowitz, "Attachment styles among young adults: a test of a four-category model." *Journal of personality and* social psychology, vol. 61, no. 2, p. 226, 1991.
- [5] M. W. BALDWIN and B. FEHR, "On the instability of attachment style ratings," *Personal Relationships*, vol. 2, no. 3, pp. 247–261, 1995. [Online]. Available: http://dx.doi.org/10.1111/j.1475-6811.1995.tb00090.x
- [6] D. Zeifman and C. Hazan, "A process model of adult attachment formation." 1997.

- [7] L. Tickle-Degnen and R. Rosenthal, "The nature of rapport and its nonverbal correlates," *Psychological Inquiry*, vol. 1, no. 4, pp. 285– 293, 1990.
- [8] L. K. Guerrero, "Attachment-style differences in intimacy and involvement: A test of the four-category model," *Communications Monographs*, vol. 63, no. 4, pp. 269–292, 1996.
- [9] M. B. Levy and K. E. Davis, "Lovestyles and attachment styles compared: Their relations to each other and to various relationship characteristics," *Journal of social and Personal Relationships*, vol. 5, no. 4, pp. 439–471, 1988.
- [10] J. A. Feeney, P. Noller, and V. J. Callan, "Attachment style, communication and satisfaction in the early years of marriage." 1994.
- [11] J. Gratch, R. Artstein, G. M. Lucas, G. Stratou, S. Scherer, A. Nazarian, R. Wood, J. Boberg, D. DeVault, S. Marsella *et al.*, "The distress analysis interview corpus of human and computer interviews." in *LREC*, 2014, pp. 3123–3128.
- [12] M. Mikulincer and O. Nachshon, "Attachment styles and patterns of self-disclosure." *Journal of Personality and Social Psychology*, vol. 61, no. 2, p. 321, 1991.
- [13] J. A. Fails and D. R. Olsen Jr, "Interactive machine learning," in Proceedings of the 8th international conference on Intelligent user interfaces. ACM, 2003, pp. 39–45.
- [14] C. Clabaugh and M. Mataric, "Exploring elicitation frequency of learning-sensitive information by a robotic tutor for interactive personalization," in *Robot and Human Interactive Communication*, 2016 RO-MAN: The 25rd IEEE International Symposium on. IEEE, 2016.
- [15] A. Tapus, M. J. Mataric, and B. Scasselati, "Socially assistive robotics [grand challenges of robotics]," *Robotics & Automation Magazine*, *IEEE*, vol. 14, no. 1, pp. 35–42, 2007.
- [16] B. Scassellati, H. Admoni, and M. Mataric, "Robots for use in autism research," *Annual review of biomedical engineering*, vol. 14, pp. 275– 294, 2012.
- [17] M. T. Schultheis and A. A. Rizzo, "The application of virtual reality technology in rehabilitation." *Rehabilitation psychology*, vol. 46, no. 3, p. 296, 2001.
- [18] J. Mumm and B. Mutlu, "Human-robot proxemics: physical and psychological distancing in human-robot interaction," in *Proceedings of the* 6th international conference on Human-robot interaction. ACM, 2011, pp. 331–338.
- [19] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *Intelligent Robots and Systems*, 2008. IROS 2008. IEEE/RSJ International Conference on. IEEE, 2008, pp. 3230–3235.
- [20] A. Tapus, C. Ţăpuş, and M. J. Matarić, "Userrobot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy," *Intelligent Service Robotics*, vol. 1, no. 2, pp. 169–183, 2008.
- [21] R. C. Fraley, M. E. Heffernan, A. M. Vicary, and C. C. Brumbaugh, "The experiences in close relationshipsrelationship structures questionnaire: A method for assessing attachment orientations across relationships." *Psychological assessment*, vol. 23, no. 3, p. 615, 2011.