Relationship Development with Humanoid Social Robots: Applying Interpersonal Theories to Human–Robot Interaction

Applicability of Interpersonal Theories of Relationship Development

The predominant paradigm for understanding interpersonal relationships is based on the concept of social exchange. The fundamental assumptions of social exchange are that people need resources to survive, other people can provide resources, and sharing and trading resources is a fundamental aspect of relationships. Although this perspective has been proposed as a valid foundation for understanding human–robot relationships, a closer examination of various theories indicates that the nature of modern HSRs may challenge their assumptions and claims.

Resources

According to the resource theory of social exchange, the resources people exchange in relationships range in how tangible or abstract they are and what function they serve. For example, money or goods are tangible economic resources, whereas love, status, and information are more abstract and social. Tangible resources are transferred from one person to another; the giver must relinquish a resource such as money or goods, and ownership shifts from the giver to the recipient. Intangible resources, in contrast, are shared between two people. Resources also vary on *particularity*, or how much a resource's value is contingent on who is providing it. For example, money spends the same whether it is received from a bank teller or a spouse, whereas love is particular and presumably more valuable coming from one's spouse rather than a bank teller.

HRI challenge

HSRs have some resources to provide and perhaps exchange with humans, particularly services and information. A robot itself, however, is a tangible good that is owned by someone, and as such does not have ownership over resources such as money or other goods and cannot transfer them to a human recipient. As HSRs are subservient to humans, they have little to offer humans in terms of status. Although it is possible for a human to form an attachment and perhaps love a robot, this is a unidirectional offering rather than a shared resource. Therefore, most social exchange resources outlined by resource theory may not be pertinent to evaluating human–robot relationships. In addition, a lack of persistent, personalized memory indicates that resources are not particularized from the robot's perspective and it is likely the human would not perceive a robot's resources as particularized either. Thus, as relational partners, robots would be perceived as interchangeable and relationships with them impersonal rather than special.

Costs, benefits, and equity

Social exchange theories also posit that humans are fundamentally self-interested and evaluate the costs and benefits they incur in a relationship. Within a relationship, individuals become interdependent through their exchange of resources, which may be more or less symmetrical. As such, relationships may be evaluated on whether these exchanges are relatively balanced, or if one partner is incurring more costs or receiving more benefits. These evaluations may be based on perceptions of equity, or whether an individual is receiving benefits or output proportional to the amount of input or costs they are incurring, particularly compared with the social norm. According to interdependence theory, one way to assess the costs and benefits of one's relationship is to compare it with other relationships, or to compare the relationship with the current partner to alternatives. If partners feel underbenefited, they are likely to experience dissatisfaction and seek to restore equity in the relationship. Over time, underbenefited partners may become dissatisfied with the relationship and terminate it, particularly if there are desirable alternatives.

HRI challenge

A major challenge to the applicability of social exchange theories is considering the nature of costs and benefits to robots. HSRs lack the motivation and desire that characterize human needs; they do not experience rewards, punishments, benefits, and costs in the ways that humans do. Furthermore, HSRs are servile to human controllers; there is a permanent inequity as they are designed to maximize benefits for humans with minimal, if any, consideration of the costs they might incur. They do not make autonomous evaluations of equity with their human controllers. They do not experience dissatisfaction, make comparisons, nor have the ability to act based on these assessments. They cannot leave their human controllers. Knowing this, humans do not have to consider their robot partner's costs or benefits, only the costs and benefits to themselves. They can make unilateral decisions without concern for the robot's wishes or well-being based on one self-serving principle: maximize my benefits and never mind the robot.

Self-disclosure

Social penetration theory (SPT) suggests that relationships develop through reciprocal self-disclosure. Individuals consider the costs and benefits of ongoing disclosure and determine whether they want to intensify the relationship.

Altman and Taylor used an onion metaphor to explain how individuals maintain many layers of self-rooted in their experiences, in which the outside layer is the publicly observable self and private information is stored in deeper layers that must be uncovered. In a developing relationship, individuals peel back these layers through reciprocal self-disclosure, proceeding through stages characterized by expanding breadth and growing depth. Breadth is characterized as the range of topics or categories that comprise the self, such as social identities, interests, and experiences. Depth involves the beliefs and values that are central to the self. According to SPT, the earliest stage of a relationship, orientation, is characterized by small talk and governed by social norms of appropriateness. In the exploratory affective stage, slightly deeper self-disclosure occurs across a broader range of topics. In the affective exchange stage, feelings of intimacy escalate as partners reveal deeper facets of the self, including values, goals, or fears. The stable exchange stage is characterized by mutual understanding, and partners are comfortable disclosing deep private matters.

HRI challenge

Given modern HSRs are constrained in their tasks and abilities, they do not have much breadth. HSRs also do not have a unique cluster of beliefs, values, and self-image that characterize depth. Although HSRs may share information, it is not based on personal experience or self-image; thus, exchanges with HSRs arguably do not qualify as self-disclosure, and they could not engage in the reciprocal self-disclosure required in a developing relationship. It should be noted that a human partner may make false attributions about the social potential of HSRs, particularly in short-term interactions. As SPT notes, social norms guide early interactions and disclosures are shallow. Humans are more likely to follow scripts of socially acceptable behavior that may be easier for HSRs to mimic, and researchers may then observe effects similar to what would be expected of a human-human interaction. Over time, however, humans would recognize an HSR's lack of personalized persistent memory, which would be necessary to build a relationship.

In summary, common HSRs are not human like enough at this time to meet the fundamental assumptions and claims of key interpersonal theories, and it is unclear when, or if, they ever will be. This challenges the popular mindset of working from the assumption that human-human findings will apply to HSRs and applying our understanding of interpersonal interactions to HRI and human–robot relationships.