Towards Understanding Trust Through Computational Cognitive Modeling

William G. Kennedy

Center for Social Simulation and Department of Computational Social Science
Krasnow Institute of Advanced Study, George Mason University
4400 University Drive, Fairfax, VA 22030, USA
wkennedy@gmu.edu

Abstract

Trust is a vital component of human society. Without trust, we would all have to continuously focus on survival, i.e., our physical security and the daily finding of food and water. With trust, we developed the division of labor that allowed us to sleep safely and develop a complex society. However, the concept of trust is very defused and not described operationally, i.e., in terms of the mechanisms by which we build, maintain, and decide to trust. There is now evidence that trust has a biological foundation [1]. That brain imaging data shows that some of the regions of our brain associated with trust are part of our high-level reasoning but others are deep within the brain. Like emotions, deep brain processes are typically fast, unconscious, and unexplainable. My hypothesis is that trust is based on cognitive mechanisms, likely the brain's slow and fast processes [2]. This talk focuses on a computational cognitive modeling approach to developing and understanding the mechanisms of trust in humans. The approach to investigating the mechanisms of trust is the use of computational cognitive modeling, specifically ACT-R [3]. Cognitive modeling is the development of declarative and procedural knowledge that "runs" on a cognitive architecture and produces outputs/behaviors comparable to those of humans. The result of this work will be an ACT-R model replicating human behavior on trust and explaining possible mechanisms of trust. However, the ACT-R system is intended to model *rational* behavior. Therefore, I will be also discussing mechanisms not normally considered to be "rational" [2]. Knowing the mechanisms associated with trust will allow a principled approach to developing intelligent systems we can and will trust.

Keywords. Cognitive modeling, trust

References

[1] F. Krueger, K. McCabe, J. Moll, N. Kreigeskorte, R. Zahn, M. Strenziok, A. Heineke, & J. Grafman, Neural Correlates of Trust, *PNAS* **104** (2007), 20084-20090.
[2] W.G. Kennedy & M.D. Bugajska, Integrating Fast and Slow Cognitive Processes. *Proceedings of the International Conference on Cognitive Modeling 2010*, 121-126.
[3] J.R. Anderson, *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press. Oxford, 2007.