Know Thyself: Self-aware Computing and Neuro-AI Strategies for Scaling Autonomous Personalization in Robotics and Healthcare

Jasmine Berry

1 Research Overview

The advent of Large Language Models such as ChatGPT and Bard sparked a call to action from a select group of researchers and brain-related AI technologists; specifically to be cautious of the rapidly-growing architecture complexity and unexpected emergent behavior in AI systems. In response, there is a rekindling interest in self-aware computing, an area of research that once laid dormant due to the difficulty in examining the epistemological nature of the mind. Socrates' maxim "To know thyself is the beginning of wisdom" underscores the importance of *understanding one's abilities and limitations*: this principle enables agents (both humans and robots) to make sense of the world through an understanding of self (Berry, 2023). A significant challenge in robotics lies in enabling robotic agents to be more contextually aware of variable situations. *Self-aware computing* refers to systems that are capable of adapting to changing environments and conditions, potentially even predicting future needs or problems, and adjusting their behavior accordingly (Kounev et al., 2017). For everyday consumer products (e.g., smartphones, productivity apps, and robotic assistive devices) the aim is to develop agents capable of providing the appropriate support in any given context. My work extends beyond *situational awareness*—concerning how an agent perceives and reacts to external stimuli and environmental factors—and instead explores the *emergence of identity and self-awareness*—focusing on how an agent's internal understanding of its purpose and role influences its actions (summarized in Figure 1). Early evidence indicates self-aware computing, developed using architectures inspired by brain-body dynamics, has fundamental implications for practical applications in precision medicine and health diagnostics (Johnson et al., 2021) and intuitive human-robot interactions (HRI) (Sheridan, 2016).

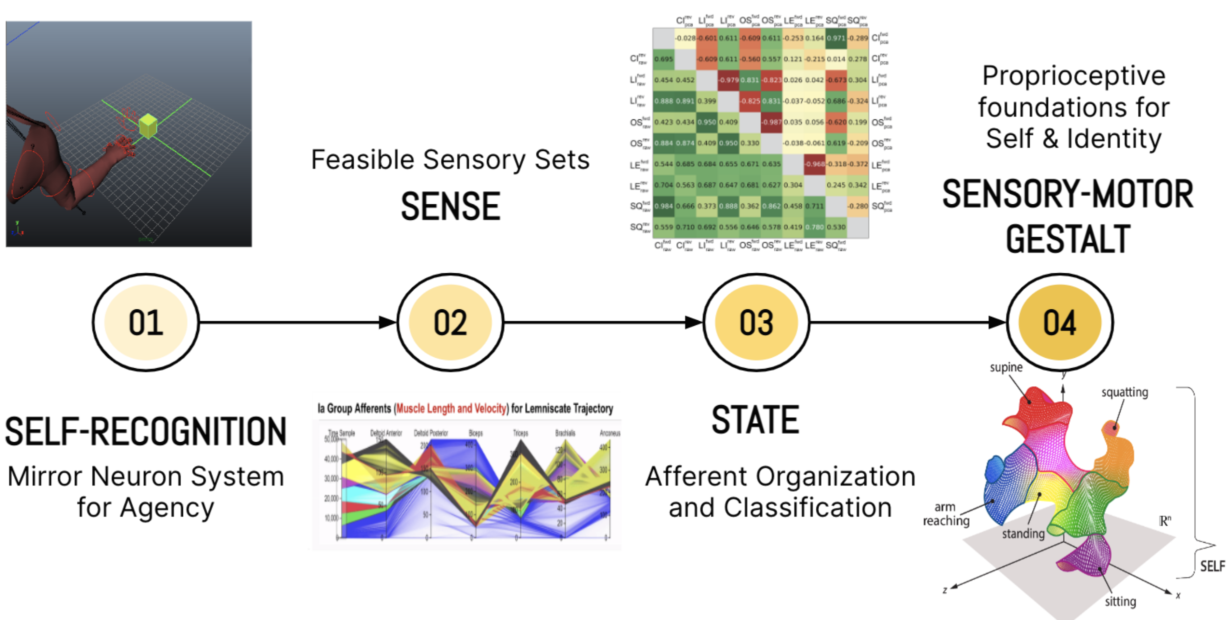

Figure 1: Research workflow for constructing body representations and self-models. In Phase 1 SELF-RECOGNITION, I identified Brain Operating Principles to extend the Mirror Neuron System 2 (MNS2) model for self-recognition and agency. In Phase 2 SENSE, I discovered the non-linear nature of proprioceptive afferents and how to explicitly analyze the high-dimensional *Feasible Sensory Sets* (FSS) that accompany limb movements. In Phase 3 STATE, I show how muscle spindles can be used for accurate state classification for identifying body movements. Lastly, in Phase 4, I proposed the *Sensory-Motor Gestalt* (smG), which is the syntactical relationship between proprioceptive and motor elements that forms a sense of identity in biological and artificial agents.

1.1 Motivation and Problem

I seek to equip agents with self-awareness and introspection to know their roles in the world while adapting their functionalities to meet the needs of the individuals (i.e., other humans and agents) they interact with in real time. To investigate this inquiry, I have addressed the following problems: 1) establishing the computational neuroscience principles for forming body schemas (Berry et al., 2023a), 2) examining active sensing from sensorimotor maps to guide movements and useful actions (Berry et al., 2017), and 3) providing computational analogues for self construction via Gestalt Laws of perception (Berry and Valero-Cuevas, 2020). I propose that self-modeling is applicable to agents that fit the definition of any system that perceives its environment through *sensors* and acts upon that environment through *effectors*. This includes software agents, intelligent agents, mobile agents, multi-agent systems, and robotic agents. Building body representations based on biological nervous systems enables one to surpass the limits of traditional robotics in multimodal integration, spatial coordination, prediction and planning, and sensory signals. However, the challenges in replicating the nervous system's mechanisms for sensorimotor mapping include sufficiently modeling the intricacies of neural plasticity, catastrophic forgetting, controlling nonlinear, over- and under-determined systems, adapting to growth, failures, tool-use, and changing environments. The core problem I am tackling involves developing systems that *autonomously* gain diagnostic knowledge of themselves (with minimal human pre-programming) to subsequently better understand and adapt to the individual needs of others.

2 Current Research and Contributions

As a graduate student at the University of Southern California, I presented a critical step in forming self-identified body schemas based on physiological simulations of proprioceptive afferent signals to determine the plausibility of biological signals' aptitude to inform the operation of a state machine or simulated mental models of the self. Building neuromorphic robots called NeuRoBots (i.e., robots that imitate the mechanisms of neural sensorimotor processing in animals) as a testbed platform was one approach to accomplishing this goal. The notion of how the anatomical brain builds a sense of self and how neuro-robotic agents can utilize body schemas (or representations) to build a sense of self have not been particularly successful due to varied and often contradictory accounts. The following sub-sections describe my contributions towards forming body schemas to meet an ongoing engineering challenge towards achieving agility, information processing, and flexibility in robotic systems (see Fig. 1).

2.1 SELF-RECOGNITION: Extending the Mirror Neuron System 2 for Grasping

In this study, I extended the Mirror Neuron System 2 (MNS2) model to explore and simulate self-recognition and self-agency in artificial agents. This extension involved creating a circuit map, comprising three key phases: Learning, Observation, and Autonomous Action/Behavior Generation. The *Learning Phase* enables the agent to integrate visual and motor body images, utilizing proprioceptive cues similar to those in the human cerebellum. The *Observation Phase* is designed to develop the ability for self-other distinction, and the final phase, *Autonomous Action/Behavior Generation*, facilitates the agent's action and adaptation based on feedback. To test and validate the extended MNS2 system, I implemented an *Agent Self-Recognition Test*, inspired by the classical Human Mirror Test. My results showed that the agent's progressively improving self-recognition capabilities in simulation can potentially parallel human cognitive development in recognition and mirroring the actions of others.

2.2 SENSE: Quantifying High-Dimensional Feasible Sensory Sets

Sensorimotor control research, both past and ongoing, has made efforts to predominately provide evidence for how the brain shapes the body. However, the counterpart to these works, how the body shapes the brain is not as extensively considered (Bernardi et al., 2015). Often not taken into account are sensory states and their effects on building the brain's body awareness which is necessary for involuntary and voluntary behavior. I investigated the flow of information underlying limb movements, and explore its significance to perceptual learning (Berry et al., 2017). I introduced the concept of trajectory-specific sensory manifolds, also called *Feasible Sensory Sets* (FSS), which are the unique multidimensional and time-varying combinations of afferent signals that obligatorily emerge during a limb movement. We use the example of *muscle spindles* (i.e., the muscle's proprioceptors for length and velocity) that arise during movements of an arm (a planar 2-DOF 6- muscle model) during the production of straight, curved and oscillatory hand movements. Through the use of parallel coordinates, we can visualize the high-dimensional evolution of the afferent signaling across muscles and tasks. I have demonstrated that limb movements give rise to a distinct sensory manifold embedded in the 12-D space of spindle information that is largely independent of the choice of muscle coordination strategy.

2.3 STATE: Organization of Sensory Afferents to Classification of Actionable States

After identifying the Feasible Sensory Sets, I have explored the classification of correlations and patterns in sensory signals associated with various gaits and states. This portion of my work resembles a form of edge computing in nature (Berry et al., 2023a). High-dimensional proprioceptive signals like those from muscle spindles are thought to enable robust estimates of bodily states. Yet, it remains unknown whether spindle signals suffice to discriminate limb movements. I found that cross-correlation of the 8D time series of raw firing rates (four Ia signals, four II signals) cannot discriminate among most movement pairs (only 29% by one measure). However, projecting these signals onto their 1st and 2nd principal components *greatly improves* discriminability of movement pairs (82% by that same measure). I concluded that high-dimensional multi-muscle proprioceptive ensembles can usefully discriminate limb states—but only after minimal pre-processing. Importantly, this may explain the documented subcortical pre-processing of afferent signals, such as cutaneous signals processing by the cat's cuneate nucleus.

2.4 SENSORY-MOTOR GESTALT: Exploring a continuum for constructing self via Gestalt Laws of perception

Limb movement produces continuous high-dimensional ensembles of afferent information that provide an *internal proprioceptive body representation* and its relationship to the environment. However, we do not yet understand how physiological proprioceptive afferents contribute to internal body representations, neuromuscular control, or even a sense of agency and self. I extended and formalized active sensing into an integrative approach—born out of a neuromechanical perspective—that sees proprioceptive and motor signals as integral parts of the same functional and perceptual continuum we call the *Sensory-Motor Gestalt* (smG) (Berry and Valero-Cuevas, 2020). The smG combines formalisms of physics, state estimation, biomechanics, differential geometry, and physiology to understand the emergence of the self in the context of proprioception and motor actions in the physical world. Proprioception, by defining *body state*, defines feasible (continuous or discrete) motor actions compatible with that state and the environment. Conversely, motor actions produce subsequent, often predictable, body states. This syntactical relationship leads to an epistemological continuum that spans body state, feasible behavior, agency, identity, and sense of self in organisms and robots.

3 Future Vision

I aim to extend my findings of machine self-representations that enable flexible behaviors towards the design of autonomous systems that integrate various levels of personalization and introspection (Berry, 2023), particularly in the realm of digital health applications. Traditional prescriptive and diagnostic healthcare models often follow a standardized approach, which falls short in addressing the unique aspects of each individual's fitness. My approach involves an *interdisciplinary* strategy, combining insights from computer science, robotics, cognitive science, psychology, neuroscience, and healthcare. I hypothesize that self-modeling and theory of mind processes can be adapted to health and HRI applications from the construction of: 1) an artificial self (i.e., simulated internal model) that emulates a neuromechanical sensorimotor body representation (Berry and Valero-Cuevas, 2020); 2) modeling edge potentials of local evidence and pairwise compatibility based on personal expectations of latent and observed variables; and 3) more computationally efficient algorithms to build from the Markov Random Fields and Factor graph models. The intellectual merit of this proposal addresses two challenges from Yang et al. (2018): 1) building an artificial theory of mind to assist with robot-human social interactions and 2) bio-inspired robots that couple sensation and action based on natural processes to maintain autonomy in novel environments.

3.1 SPACE and SHAPE for Emergent Self Models

Once the smG is identified then I intend to incorporate approaches of 1) SPACE which focuses on the aggregation and collection of all experienced states of the action phenomenon while exploit inter-state similarities and derivations and 2) STATE which forms topologies and geometry of evolving sensorimotor manifold shapes. I seek to leverage topological and geometric methodologies to develop personalization models in robotics. By constructing high-dimensional manifolds that represent diverse sensorimotor states, robotic interactions can be tailored to individual preferences and behaviors. Utilizing topological data analysis, I identify distinct patterns within the sensorimotor data, essential for customizing robotic responses. Additionally, differential geometry aids in understanding the complex relationships between sensorimotor states, crucial for creating adaptive algorithms that respond to individual user differences. Integrating these techniques into neural network designs, I am advancing the development of robotic systems that not only respond appropriately but also offer personalized experiences, enhancing the scope of robotics in applications like personalized assistance and customized therapeutic interventions. My work will pioneer a new direction in blending geometric data analysis of manifolds with practical robotics applications, aiming to create more responsive and user-tailored assistive robotic systems.

3.2 Neuro-cognitive Models of Identity for Theory of Mind Applications

As part of my postdoctoral research at the University of Michigan in collaboration with Prof. Chad Jenkins, I then focused on the need to study identity in multi-agent task allocations and ally vs. adversary teaming (Berry et al., 2023b). Advancements in autonomous agents have led to an increasingly ubiquitous presence of robots in human environments where social and physical interaction is expected. Such environments are often composed of heterogeneous agents with disparate action capabilities, intentions, and motivations. Intra- and inter-agent dissimilarity often prevents enacting effective behavioral skills (e.g., collaboration, communication, coordination) towards dynamic task allocation objectives. I proposed a Bayesian probabilistic inference approach, Multi-Robot Belief Propagation with Identity constraints (MRPB-I), for 1) decentralized task allocation in multi-agent systems and 2) modeling task affinity using personal identities. MRPB-I leverages competing costs of individual and group capabilities that result in less error-prone convergence to steady state, scalability without loss of accuracy, and sensitivity to environmental dynamics. I presented an implementation of MRBP-I as a distributed algorithm that weighs factors of both individual and cooperative perception in an energy-minimizing task allocation scheme.

3.3 Using the Neuroscience of Self-Awareness and Consciousness to Improve Personalized Patient Care and Human Robot Interactions

How can neural models of emergent cognitive properties associated with self-awareness and consciousness help advance AI systems, which could then simultaneously enable better 1) contextually-aware robotic agents and 2) predictive applications for precision health and medicine? My research methodology is to develop cognitive computing models of identity that investigate the neural correlates of how individuals perceive themselves and their surroundings. Such models have demonstrable use-cases in fostering self-organizing architectures for autonomous systems that can adapt and reorganize their internal structures in response to external stimuli and unpredictable changes; particularly in scenarios where pre-programming constraints at design time are insufficient. Precision health has not been fully solved due to the complexity of modeling human health, need for *explainable decision making*, and the degree of personalization required to be processed by the ever-growing number of personal medical devices (Bender et al., 2022; Bender and Berry, 2023). Success in this research endeavor hinges on not just technological innovation but also on navigating complex ethical, social, accessibility, equability, and practical challenges for creating a society that optimizes towards individual behavioral and lifestyle factors.

References

Bender, B. F. and Berry, J. A. (2023). Trends in Passive IoT Biomarker Monitoring and Machine Learning for Cardiovascular Disease Management in the US Elderly Population. *Advances in Geriatric Medicine and Research*, 5(1).

Bender, B. F., Johnson, N. J., Berry, J. A., Frazier, K. M., and Bender, M. B. (2022). Automated Urinal-Based Specific Gravity Measurement Device for Real-Time Hydration Monitoring in Male Athletes. *Frontiers in Sports and Active Living*, 4:921418.

Bernardi, N. F., Darainy, M., and Ostry, D. J. (2015). Somatosensory contribution to the initial stages of human motor learning. *Journal of neuroscience*, 35(42):14316–14326.

Berry, J. A. (2023). Agent assessment of others through the lens of self. Accepted for Presentation at AAAI-FSS.

Berry, J. A., Marjaninejad, A., and Valero-Cuevas, F. J. (2023a). Edge Computing in Nature: Minimal pre-processing of multi-muscle ensembles of spindle signals improves discriminability of limb movements. *Frontiers in Physiology*, 14.

Berry, J. A., Olson, E. A., Gilbert, A., and Jenkins, O. C. (2023b). A Case of Identity: Enacting Robot Identity with Belief Propagation for Decentralized Multi-Agent Task Allocation. *IEEE RO-MAN*.

Berry, J. A., Ritter, R., Nagamori, A., and Valero-Cuevas, F. J. (2017). The neural control of movement must contend with trajectory-specific and nonlinearly distorted manifolds of afferent muscle spindle activity. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 1188–1194. IEEE.

Berry, J. A. and Valero-Cuevas, F. J. (2020). Sensory-Motor Gestalt: Sensation and Action as the Foundations of Identity, Agency, and Self. In *ALIFE 2020: The 2020 Conference on Artificial Life*, pages 130–138. MIT Press.

Johnson, K. B., Wei, W.-Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., and Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and translational science*, 14(1):86–93.

Kounev, S., Lewis, P., Bellman, K. L., Bencomo, N., Camara, J., Diaconescu, A., Esterle, L., Geihs, K., Giese, H., Götz, S., et al. (2017). The notion of self-aware computing. *Self-Aware Computing Systems*, pages 3–16.

Sheridan, T. B. (2016). Human–robot interaction: status and challenges. *Human Factors*, 58(4):525–532.

Yang, G.-Z., Bellingham, J., Dupont, P. E., Fischer, P., Floridi, L., Full, R., Jacobstein, N., Kumar, V., McNutt, M., Merrifield, R., et al. (2018). The grand challenges of science robotics. *Science robotics*, 3(14):eaar7650.