Research Outline:
Machine Autonomy and Developmental Autonomous Behavior

Updated on Aug 7, 2025 | Created on November 16, 2019

Satoru Isaka

ORCID: 0000-0001-5670-0418

sisaka@visiondelmar.com

This document outlines my research objective, approach, and key results on machine autonomy with updated descriptions of current efforts.

1. Research objective and scope

The core objective of this research is to derive and apply a principled and practical framework to understand the developmental autonomous behavior (DAB) of machines. DAB refers to "*the general ability of a machine to acquire new skills and behavior from its birth to maturity on its own without human intervention*" [1]. It draws inspiration from ethology, the study of animal behavior, to derive ultimate-proximate explanations of the causality, mechanisms, and circumstances of behavior emergence grounded in the physical interactions and sensory experiences of machines in their environment.

The key difference between this research and related studies, such as artificial intelligence, cognitive science, and robotics, lies in its objective and perspective. Unlike others, this research does not aim to understand human cognition or replicate human capabilities. Instead, it strictly focuses on identifying the systems and processes that achieve autonomy from the perspective of machines. Because of this distinction, this research is *free from* the obligation to justify theories in terms of anthropocentric, ill-defined, intangible, and subjective concepts such as beliefs, consciousness, intelligence, and sentience. By eliminating these concepts from its scope, this research is *free to* close the loop between deductive and inductive analyses of machine behavior with direct access to the machines' internals, which is a major obstacle in human studies.

Granting autonomy to machines is a choice by humans, yet its implications are far more profound than most realize. This research directly addresses this important yet underappreciated issue in integrating autonomy into automation, and aims to deepen our understanding of machine autonomy. I believe that this is paramount to the safety of our society and thus a critical area of exploration for the advancement of automation science and engineering.

2. Approach

The concept of autonomy is central to this research. Based on the principles of embodiment and situatedness, the proposed framework defines autonomy in the ecological sense of survival and operation continuity. Machine autonomy is, therefore, tested in terms of the ability to maintain its operation in perpetuity, including the ability to find a way to supply its own energy without human intervention.

The framework posits that value systems for the purpose of autonomy drive the successive development of memory functions, resulting in progressive changes in behavior from innate reflexive to episodic, procedural, and autonomic behavior. The framework also derives a logical explanation for the transformation process that enables a physical sensorimotor system to become a symbol-like concept processor, fostering conceptual and social behavior development. By precisely defining the concept of autonomy and the role of value systems, the framework provides the principles of developmental autonomous behavior emerging from experience and learning.

The framework is structured for utility to build and analyze machines that exhibit developmental autonomous behavior by incorporating two practical tools: machine learning (ML) and cyber-physical systems (CPS), but in a unique way. Moving beyond traditional machine learning systems, this research examines how autonomy enables a robot to actively and intentionally shape its behavior based on its own value system, rather than statistically optimizing a performance measure given by humans. This approach broadens the role of ML algorithms from data approximators to self-learning systems by exploiting their algorithmic attributes, thereby contributing to various aspects of behavior emergence in autonomous systems. It also provides a general architecture that leverages the distributed computing and communication capabilities of CPS, necessary for continuous autonomous operation in practical applications.

3. Key Results

1. Derived and documented the fundamental principles of DAB in machines in [1]. The article provided extensive coverage and explanations of the backbone concepts, tracing historical progress, surveying literature, identifying core challenges and issues, and precisely defining terminology. Inspired by biological views of behavioral causation, the article emphasized principled explanations not only to the "how" question on mechanisms but also to the "why" question on causation of behavior development.
2. Demonstrated the DAB framework in a minimally configured mobile robot (documented in [2][6]).
3. Identified the key concepts on the taxonomy of autonomy and ML algorithms for behavior emergence (documented in [3]).
4. Developed a general taxonomic classification and identification system for robots and autonomous systems based on the DAB principles (documented in [4][5]).
5. Completed the first part of a general philosophical framework of machine autonomy (documented in [7]).
6. Wrote a perspective piece that critiques emerging linguistic centralism (documented in [8]).
7. Shared my thoughts on self-directed interpretation and kinematic expression in autonomous robots with UCSD faculties (documented in [9][10]).

4. Current Topics

The figure below shows the current efforts and expected outputs within the overall research framework.

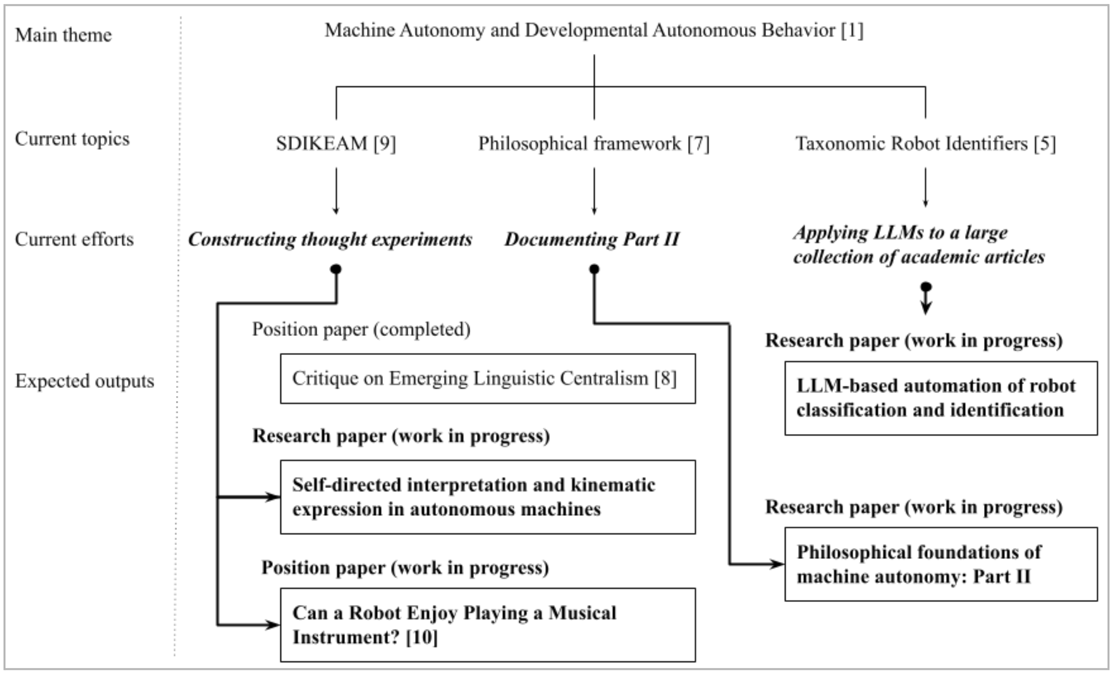

Figure. Current research efforts highlighting the topic relationships and expected outputs

1. Self-directed interpretation and kinematic expression in autonomous machines (SDIKEAM)

Following the key result 3.7, I am working on constructing a series of thought experiments.

Unlike traditional computational models for images and language, music offers a unique framework for interpretation, expression, motion, and interaction. It challenges machines to engage in self-directed interpretation and kinematic expression, rather than predetermined execution. This exploration begins with a series of thought experiments with a simple robot, consisting of a sensor unit that detects environmental signals and a computing unit that processes inputs to generate outputs, controlling the movement of mechanical arms that produce percussive sounds. Using this hypothetical machine, and by progressively adding physical components and computational elements, we ask:

1. *Why and how can the robot construct its own interpretations of phenomenal experience?*
2. *Why and how can the robot refine its motions to express a recognizable behavioral signature?*
3. *Why and how can the robot control its behavior to create collaborative sound with external sources?*
4. *What are the implications of autonomous systems capable of achieving 1, 2, and 3?*

By bridging conceptual examination with physical bodies, this research seeks to uncover the nature of machine autonomy through music and embodied learning, leading us to a better understanding of what it means to grant autonomy to machines.

2. Philosophical framework of machine autonomy

Following the key result 3.5, I am working on the second part.

This topic explores the fundamental yet often overlooked question: *what does it mean to grant autonomy to machines*? It explores the subject through four essential queries: what can machines know? (question of knowledge), what can machines do? (question of behavior), what constitutes autonomy? (question of existence), and what does it mean for machines to be autonomous? (question of freedom). By presenting philosophical underpinnings with logical answers to these queries, it provides critical insights into the essence of machine autonomy. It argues that the lack of fundamental understanding, rather than regulatory or technical deficits, poses the greatest challenge to responsibly integrating autonomous systems into our society.

3. LLM-based robot identifier

I am applying the key result 3.4 to a collection of articles by using large-language models.

This topic is a practical application of the taxonomic robot identifiers described in [4][5]. By integrating a large-language model via API with curated prompts, the system assigns a taxonomic identification code to a document that describes autonomous systems. For example, an academic conference proceeding that contains hundreds of articles can be automatically analyzed with statistics on the types of papers presented.

List of Publications and Documents Referenced

[1] S. Isaka, "Developmental Autonomous Behavior: An Ethological Perspective to Understanding Machines," in *IEEE Access*, vol. 11, pp. 17375-17423, 2023, doi: 10.1109/ACCESS.2023.3246840.

[2] S. Isaka, "An Ethological Analysis of Developmental Behavior in Machines," in *Proc. 2023 IEEE International Conference on Development and Learning (ICDL)*, Macau, China, 2023, pp. 79-86, doi: 10.1109/ICDL55364.2023.10364472.

[3] S. Isaka, "Autonomy in Cognitive Development of Robots: Embracing Emergent and Predefined Knowledge and Behavior," in *Proc. 2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*, Bari, Italy, 2024, pp. 1353-1360, doi: 10.1109/CASE59546.2024.10711540.

[4] S. Isaka, "A Taxonomic Classification and Identification System for Robots: Abstract," in *Proc. 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Kuching, Malaysia, 2024, pp. 3799-3800, doi: 10.1109/SMC54092.2024.10831650.

[5] S. Isaka, "Taxonomic Robot Identifiers: Toward General Classification and Oversight for Autonomous Systems," in *IEEE Access*, vol. 13, pp. 101801-101816, 2025, doi: 10.1109/ACCESS.2025.3578870.

[6] S. Isaka, “Behavior-Based Robotics, Cyber-Physical Systems, and Machine Learning: United as a Practical Pathway Toward Machine Autonomy,” unpublished.

[7] S. Isaka, "Philosophical Foundations of Machine Autonomy: What It Means To Grant Autonomy To Machines Part 1: Introduction and First Query," *Preprint in ResearchGate*, May 9, 2025, Available: https://doi.org/10.13140/RG.2.2.10280.48643.

[8] S. Isaka, “Critique on Emerging Linguistic Centralism: How Language Distorts Our Understanding of AI and Humanity,” *Submitted for publication*, August 5, 2025.

[9] S. Isaka, “Self-directed interpretation and kinematic expression in autonomous robots: an interdisciplinary exploration in philosophy, music, mathematics, and engineering,” *Internal report*, May 31, 2025.

[10] S. Isaka, “Can a Robot Enjoy Playing a Musical Instrument?” *Internal report*, June 16, 2025.