

# Object2Plan: An Ontological Approach to Automated Generation of Assembly Plans from Objects

Jona Thai<sup>1,\*</sup>, Michael Grüninger<sup>1</sup>

<sup>1</sup>Department of Mechanical and Industrial Engineering, University of Toronto, Ontario, Canada M5S 3G8

## Abstract

Robotic object assembly remains a core challenge in automated planning, demanding precise reasoning about both object structure and task execution. Existing symbolic approaches (e.g., PDDL) and data-driven methods (e.g., LLMs) each fall short: the former lack scalability, while the latter lack semantic grounding. Despite the central role of parthood and connection, no formal mereological foundation currently unifies object structure and assembly planning. This paper introduces a mereological framework for object assembly based on CISCO, a non-classical mereology for connected induced substructures. We define parallel mereologies for object components and assembly (sub)activities and prove them order-isomorphic. Within this framework, all possible assembly plans can be embedded within a “master plan,” linking the semantics of physical composition with those of task execution and enabling more principled reasoning over the space of possible assembly plans.

## Keywords

robotic assembly, mereotopology, robotics, task planning

## 1. Introduction

Object assembly is a deceptively simple task. Even something as trivial as joining Part A with Part B can become impossible without an understanding of the parthood and connection relations at play. These relations underpin not only the physical structure of objects but also the logic of the actions required to assemble them. Consequently, a substantial body of research has focused on developing reliable robotic assembly systems, which are central to applications ranging from high-mix manufacturing to consumer service robotics.

Two dominant paradigms currently exist in robotic assembly planning: symbolic, rule-based approaches (e.g., the Planning Domain Definition Language, PDDL) and data-driven machine learning methods (e.g., large language models, neural nets)[1][2]. The former offers reliability and interpretability but becomes brittle with scale. The latter exhibits impressive generalization ability but can be unreliable when faced with statistical long-tail events. Both paradigms, however, encounter significant limitations when addressing complex assemblies—those involving numerous subassemblies, intricate contact relations, and combinatorially large task spaces. As a result, current research efforts often seek hybrid strategies, combining symbolic representations with data-driven inference to balance generalization with rigor.

Despite these advances, several aspects of assembly planning remain surprisingly underexplored. For one, despite the undeniable role of parthood and connection plays in robotic assembly, there are no explicit axiomatizations of mereology in the literature. Without formal axioms as a backbone, data structures can only capture the taxonomy of an assembly plan, not its semantics. Another overlooked area is the characterization of the relationship between an object and its assembly plan, and by extension the relationship among all possible assembly plans for a given object. Existing approaches typically focus on generating a single feasible or optimal plan, but seldom on the structural space of possible plans as a whole. By mathematically defining this space of of potential assembly plans, these insights can be used to develop better algorithms to constrain the search space.

This paper addresses the above-listed gaps by introducing a formal mereological framework that unifies the representation of objects, task plans, and their occurrences. We define a mereology of

*RobOntics'25: International Workshop on Ontologies for Autonomous Robotics, December 10, 2025*

\*Corresponding author.

✉ jona.thai@mail.utoronto.ca (J. Thai); gruninger@mie.utoronto.ca (M. Grüninger)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

object components in parallel with a mereology of assembly (sub)activities with CISCO, a non-classical mereology for connected induced substructures [3]. We prove that the object component mereology is order isomorphic to the mereology of assembly activities. Within this framework, individual assembly plan occurrences(i.e., specific sequences or realizations of assembly actions) are surjective to the mereology of assembly activities. Put simply, all possible task plan sequences can be effectively “embedded” within a master task plan. In doing so, we connect the semantics of physical composition with the semantics of task execution, offering a unified foundation for reasoning about assembly at both the object and activity levels.

## 2. Related Work

Regardless of whether the approach is symbolic or learning-based, representation lies at the core of robotic assembly planning. Accordingly, we provide a brief overview of the current state of the art in object and task plan representations within this domain.

### 2.1. Object Representation

Given the three-dimensional nature of most assembly objects, the most common representation is through 3D computer-aided design (CAD) files. Two predominant forms of 3D object representation within CAD systems are constructive solid geometry (CSG) and boundary representation (B-rep)[4][5]. CSG models objects as compositions of primitive solids (e.g., cubes, cylinders, spheres) combined through Boolean operations such as union, difference, and intersection. In contrast, B-rep defines an object through its boundary surfaces—vertices, edges, and faces—enabling precise surface-level modeling suitable for visualization, manufacturing, and simulation.

Due to its higher fidelity, B-rep is currently the dominant CAD form. One of the most notable 3D model datasets that utilizes B-rep is the Autodesk Fusion360 Assembly Dataset, which provides a large corpus of real-world CAD assemblies with rich parametric and hierarchical information [6]. Of the many data formats available, of particular interest is the assembly graph model - where the vertices of the graph are components and the graph edges denote parent-child relationships. This allows it to capture the relationship between components and subcomponents through a parthood hierarchy [7].

Other datasets following similar paradigms include the NVIDIA Omniverse Automate Dataset, which focuses on industrial assembly scenes with annotated kinematic and physical properties, and even the Assembly101 Dataset, which provides RGB-D video sequences paired with 3D object models for studying human-robot collaborative assembly [8] [9].

It is interesting to note that at its core, all of these different approaches seek to model parthood and connection within objects, meaning these relationships and properties are most core to robot assembly activity.

### 2.2. Task Plan Representation

A critical challenge in robotic assembly lies in representing task-level plans in a manner that enables efficient search by automated planners while maintaining sufficient geometric and kinematic specificity to support both generalization across tasks and formal verification. The representation must strike a delicate balance: overly abstract formulations hinder grounding in the physical domain, whereas overly detailed encodings rapidly become intractable for symbolic planners.

The lingua franca of symbolic task planning has long been declarative languages such as the Planning Domain Definition Language (PDDL) and its predecessor, the Stanford Research Institute Problem Solver (STRIPS)[10][11]. In these frameworks, a planning problem is defined in terms of objects, initial states, and goal states, while the domain specifies the available actions and predicates. Despite their expressiveness and clarity, such definitions are often ad hoc and fail to scale gracefully as the complexity of the solution space increases. The brittleness of purely symbolic formulations has therefore motivated extensive research into more robust or hybrid representations.

One notable direction, particularly relevant to robotic assembly, is Assembly Sequence Planning (ASP). ASP addresses the problem of determining a feasible and efficient order for assembling multiple components, often represented as a directed graph of assembly operations. Precedence constraints—derived from geometric, physical, or functional dependencies between parts—govern the generation of these sequences. Classical ASP approaches leverage heuristic search and constraint reasoning, while more recent work integrates geometric reasoning and physical feasibility analysis to ensure executable plans.

An alternative paradigm, assembly-by-disassembly, adopts a physics-based simulation perspective [2]. Instead of symbolically defining assembly rules, it simulates disassembly actions to infer feasible assembly orders. With the availability of large-scale datasets capturing realistic assembly and disassembly trajectories, machine learning models—including deep neural networks—can generalize across a wide range of geometries and joint types. While this approach offers flexibility and data-driven adaptability, it continues to face limitations in modeling complex assemblies involving multiple contact types, hierarchical structures, and intricate connection relations.

More recently, neurosymbolic approaches have emerged as a promising middle ground. These methods aim to combine the interpretability and reliability of symbolic task planning with the adaptability and perception robustness of learning-based systems. By embedding symbolic constraints within learned representations (like ontologies!)—or conversely, grounding symbolic plans in continuous perception and action spaces—neurosymbolic frameworks strive to achieve both formal structure and empirical flexibility [12] [13][1]. In the context of robotic assembly, such approaches hold potential for unifying geometric reasoning, task-level planning, and sensory feedback within a coherent and semantically grounded framework.

### 2.3. Object-Task Plan Representation

Looking at the literature, we can conclude that accurately representing parthood, connection and hierarchy are tantamount to robotic task assembly. This rings true across all these diverse methods. Yet, none of these approaches explicitly utilize mereology - variables are captured through data structures, but semantics remain unclear. Relations suggestive of parthood or connectivity (e.g., part-of, attached-to, contained-in) are often included, yet lack correspondence to a formal axiomatization. Formal specification allows for models to be characterized, increasing reliability and explainability. The goal of this paper is to address that gap.

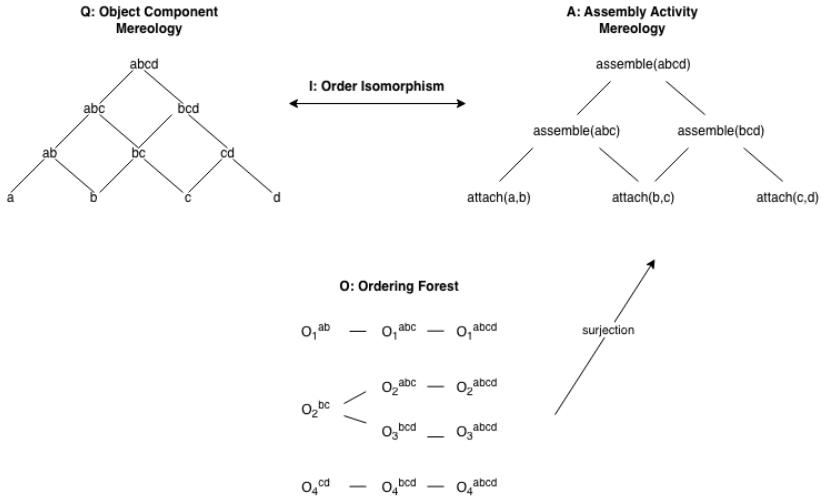
## 3. Object2Plan

The key contribution of this paper is the formalization of the relationship between the parts of an object and the parts(subactivities) of a task plan. Our approach is best illustrated by the diagram in Figure 1. The primary idea is that the *mereology of connected components of the object is isomorphic to the mereology of activities that assemble the object*. Moreover, the mereology of assembly activities describes the set of all possible occurrences of an assembly plan for the object.

### 3.1. Mereologies for Objects and Assembly Activities

Classical mereology is based on the assumption that any two overlapping elements have a sum, yet when we consider the problem of object assembly, this assumption is not valid. Instead, we find that mereological sums must always be connected objects.

We also observe that existing approaches to assembly planning represent an object as a simple graph; in this case, the components of the object should correspond to connected induced subgraphs of that graph. The work in [3] introduced the parthood and connection structure on the set of connected induced subgraphs of a graph and used this as the basis of the representation theorem for a new mereotopology,  $T_{cisco\_mt}$ , in which the sum of two elements exists iff they are connected.



**Figure 1:** Object2Plan Relationship Diagram

We therefore propose two distinct mereologies that are each logically synonymous to  $T_{cisco\_mt}$ , but with different signatures. The mereology on physical objects,  $T_{object\_component}$ , has the signature **{object(x), componentOf(x,y), comp\_sum(x,y,z), comp\_overlaps(x,y), comp\_covers(x,y)}**:

**Definition 1.**  $T_{object\_component}$  is the following sentences:

$$\forall x, y \text{ componentOf}(x, y) \supset \text{object}(x) \wedge \text{object}(y) \quad (1)$$

$$\forall x \text{ object}(x) \supset \text{componentOf}(x, x) \quad (2)$$

$$\forall x, y \text{ componentOf}(x, y) \wedge \text{componentOf}(y, x) \supset (x = y) \quad (3)$$

$$\forall x, y, z \text{ componentOf}(x, y) \wedge \text{componentOf}(y, z) \supset \text{componentOf}(x, z) \quad (4)$$

$$\forall x, y \text{ comp\_covers}(x, y) \supset \exists z \text{ atom}(z) \wedge \text{comp\_sum}(z, y, x) \wedge \neg \text{componentOf}(z, y) \quad (5)$$

$$\forall x, y \text{ comp\_overlaps}(x, y) \supset \exists z \text{ comp\_sum}(x, y, z) \quad (6)$$

$$\forall x, y \neg \text{componentOf}(x, y) \supset \exists z \text{ componentOf}(z, x) \wedge \neg \text{comp\_overlaps}(z, y) \quad (7)$$

$T_{activity\_mereology}$ , which is the mereology on assembly activities, has the signature **{assembly\_activity(x), subassembly(x,y), assembly\_covers(x,y), assembly\_overlaps(x,y), assembly\_sum(x,y)}**:

**Definition 2.**  $T_{assembly\_activity}$  is the following set of sentences:

$$\forall x, y \text{ subassembly}(x, y) \supset \text{assembly\_activity}(x) \wedge \text{assembly\_activity}(y) \quad (8)$$

$$\forall x \text{ assembly\_activity}(x) \supset \text{subassembly}(x, x) \quad (9)$$

$$\forall x, y \text{ subassembly}(x, y) \wedge \text{subassembly}(y, x) \supset (x = y) \quad (10)$$

$$\forall x, y, z \text{ subassembly}(x, y) \wedge \text{subassembly}(y, z) \supset \text{subassembly}(x, z) \quad (11)$$

$$\forall x, y \text{ assembly\_covers}(x, y) \supset \exists z \text{ atom}(z) \wedge \text{assembly\_sum}(z, y, x) \wedge \neg \text{subassembly}(z, y) \quad (12)$$

$$\forall x, y \text{ assembly\_overlaps}(x, y) \supset \exists z \text{ assembly\_sum}(x, y, z) \quad (13)$$

$$\forall x, y \neg \text{subassembly}(x, y) \supset \exists z \text{ subassembly}(z, x) \wedge \neg \text{assembly\_overlaps}(z, y) \quad (14)$$

### 3.2. Multigeometry

To formally represent the isomorphism between the mereology of object components  $\mathbb{O}$  and the mereology of assembly subactivities  $\mathbb{A}$  and their occurrences  $\mathbb{O}$ , we consider a class of structures known as multigeometries, which represent homomorphisms between mereologies.

**Definition 3.**  $\mathbb{Q} \oplus \mathbb{I} \oplus \mathbb{A}$  is an assembly multigeometry iff

1.  $\mathbb{Q} = \langle P, \preccurlyeq \rangle$  such that  $\mathbb{Q} \in \mathfrak{M}^{\text{cisco}}$ ;
2.  $\mathbb{A} = \langle L, \leq \rangle$  such that  $\mathbb{A} \in \mathfrak{M}^{\text{cisco}}$
3.  $\mathbb{I} = \langle P, L, I \rangle$  such that  $\mathbb{I} \in \mathfrak{M}^{\text{bipartite}}$
4.  $N^{\mathbb{I}}(U^{\mathbb{Q}}(\mathbf{x})) \subseteq U^{\mathbb{A}}(N^{\mathbb{I}}(\mathbf{x}))$

We denote the class of assembly multigeometries by  $\mathfrak{M}^{\text{assembly\_plan}}$ .

By conditions 1 and 2, there exists two partial orderings  $\mathbb{Q}$  and  $\mathbb{A}$ . Condition 3 guarantees that the incidence relation  $\mathbb{I}$  represents a mapping  $\mu : \mathbb{Q} \rightarrow \mathbb{P}$ . Condition 4 guarantees that this mapping is an order isomorphism.

**Theorem 1.** Let  $\text{Iso}(\mathbb{Q}, \mathbb{A})$  denote the set of all poset order isomorphisms between the partial orderings  $\mathbb{Q}, \mathbb{A}$ . If

$$\mathcal{P} = \{ \mu : \mu \in \text{Iso}(\mathbb{Q}, \mathbb{A}) \in \mathbb{Q}, A \in \mathfrak{M}^{\text{cisco}} \}$$

then there is a bijection  $\varphi : \mathcal{P} \rightarrow \mathfrak{M}^{\text{assembly\_plan}}$  such that

$$\varphi(\mu) = \mathbb{Q} \oplus \mathbb{I} \oplus \mathbb{A}$$

iff  $\mu : \mathbb{Q} \rightarrow \mathbb{A}$  and

$$\mu(x) = N^{\mathbb{I}}(x) \cap L$$

This guarantees the property that each assembly multigeometry  $\mathbb{Q} \oplus \mathbb{I} \oplus \mathbb{A}$  corresponds to an order isomorphism that maps the object component mereology  $\mathbb{Q}$  to the assembly activity mereology  $\mathbb{A}$ . The incidence structure  $\mathbb{I}$  represents this order isomorphism in the sense that the subordering induced by the non-atoms of the object component mereology are isomorphic to the atoms of the subactivity mereology.

Combining all of the axioms together gives us the complete ontology for Object2Plan:

**Definition 4.**  $T_{\text{assembly\_plan}}$  is the extension of  $T_{\text{object\_component}} \cup T_{\text{assembly\_activity}}$  with the following set of sentences:

$$(\forall x, a) \text{assemble}(x, a) \supset \text{object}(x) \wedge \text{assembly\_activity}(a) \quad (15)$$

$$(\forall p) \text{object}(p) \supset \neg \text{assembly\_activity}(p). \quad (16)$$

$$(\forall l_1, l_2, p) \text{assembly\_activity}(l_1) \wedge \text{assembly\_activity}(l_2) \wedge \text{object}(p) \\ \wedge \text{assemble}(p, l_1) \wedge \text{assemble}(p, l_2) \supset (l_1 = l_2). \quad (17)$$

$$(\forall p_1, p_2, l) \text{assembly\_activity}(l) \wedge \text{object}(p_1) \wedge \text{object}(p_2) \\ \wedge \text{assemble}(p_1, l) \wedge \text{assemble}(p_2, l) \supset (p_1 = p_2). \quad (18)$$

$$(\forall x, y, l_1, l_2) \text{componentOf}(x, y) \wedge \text{assembly\_activity}(l_1) \wedge \text{assembly\_activity}(l_2) \\ \wedge \text{assemble}(x, l_1) \wedge \text{assemble}(y, l_2) \supset \text{subassembly}(l_1, l_2). \quad (19)$$

$$(\forall x) (\text{object}(x) \supset (\exists y) (\text{assembly\_activity}(y) \wedge \text{assemble}(x, y))). \quad (20)$$

$$(\forall x, y, l_1, l_2) \text{assemble}(x, l_1) \wedge \text{assemble}(y, l_2) \wedge \text{object}(x) \wedge \text{object}(y) \\ \wedge \text{subassembly}(l_1, l_2) \supset \text{componentOf}(x, y). \quad (21)$$

$$(\forall l) (\text{assembly\_activity}(l) \supset (\exists p) (\text{object}(p) \wedge \text{assemble}(p, l))). \quad (22)$$

These axioms specify the mereologies on both the objects being assembled and also the activities that assemble the objects. The guarantee that these two mereologies are isomorphic provides a way of evaluating the correctness of assembly plans with respect to the mereological structure of the object.

### **3.3. Ontological Consequences**

Identity criteria for objects cannot be reduced merely to their primitive components, since two distinct objects may be composed of the same basic elements yet remain different due to the ways those elements are connected. What distinguishes them is not the underlying parts, but the structure of their composition—the mereological relations that hold among those parts. In this sense, even if a mapping or even an isomorphism exists between their components, the objects remain distinct because their identities are determined by their composition. This distinction becomes especially important when contrasting well-defined classes of activity with non-deterministic activities, since the criteria for identity must account not only for the presence of components but also for the specific structural and relational configurations that constitute the object.

## **4. Discussion & Future Work**

This paper provides a formal mereological framework that unifies the representation of objects, task plans, and their occurrences. Specifically, a mereology for object components and a mereology for assembly activities are defined utilizing CISCO, a non-classical mereology for connected induced substructures. We proved that the object mereology is order isomorphic to the assembly activity mereology, where the subordering induced by the non-atoms of the object component mereology are isomorphic to the atoms of the subactivity mereology. For future work, we intend to extend this work to cover a wider range of parthood and connection relation types, and further utilize our insights to guide development of robotic assembly algorithms. In particular, we intend to define the relationship between assembly plan occurrences and the assembly plan mereology, with the requirement that there be a 1-1 correspondence between occurrences of the plan and maximal chains in the mereology of assembly activities.

## **5. Declaration on Generative AI**

During the preparation of this work, the author(s) used ChatGPT in order to: Paraphrase and reword, translation of APA-style citations into BibTex format. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

## References

- [1] M. Kwon, Y. Kim, Y. J. Kim, Fast and accurate task planning using neuro-symbolic language models and multi-level goal decomposition, in: IEEE International Conference on Robotics and Automation (ICRA), 2025.
- [2] Y. Tian, J. Xu, Y. Li, J. Luo, S. Sueda, H. Li, K. D. Willis, W. Matusik, Assemble them all: Physics-based planning for generalizable assembly by disassembly, ACM Transactions on Graphics 41 (2022).
- [3] M. Grüninger, C. Chui, Y. Ru, J. Thai, A mereology for connected structures, in: Formal Ontology in Information Systems Proceedings of the 11th International Conference, (FOIS 2020), volume 330 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2020, pp. 171–185. doi:10.3233/FAIA200670.
- [4] A. A. G. Requicha, H. B. Voelcker, Constructive solid geometry, Technical Report TM-25, University of Rochester, Production Automation Project, 1977.
- [5] B. G. Baumgart, A polyhedron representation for computer vision, in: Proceedings of the National Computer Conference, 1975, pp. 589–596.
- [6] K. D. D. Willis, Y. Pu, J. Luo, H. Chu, T. Du, J. G. Lambourne, A. Solar-Lezama, W. Matusik, Fusion 360 gallery: A dataset and environment for programmatic cad construction from human design sequences, ACM Transactions on Graphics (TOG) 40 (2021).
- [7] K. D. Willis, P. K. Jayaraman, H. Chu, Y. Tian, Y. Li, D. Grandi, A. Sanghi, L. Tran, J. G. Lambourne, A. Solar-Lezama, W. Matusik, Joinable: Learning bottom-up assembly of parametric cad joints, arXiv preprint arXiv:2111.12772 (2021).
- [8] B. Tang, I. Akinola, J. Xu, B. Wen, A. Handa, K. V. Wyk, D. Fox, G. S. Sukhatme, F. Ramos, Y. Narang, Automate: Specialist and generalist assembly policies over diverse geometries, in: Robotics: Science and Systems (RSS), 2024.
- [9] F. Sener, D. Chatterjee, D. Sheleporov, K. He, D. Singhania, R. Wang, A. Yao, Assembly101: A large-scale multi-view video dataset for understanding procedural activities, CVPR (2022).
- [10] D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, D. Wilkins, PDDL—The Planning Domain Definition Language, Version 1.2, Technical Report, Yale Center for Computational Vision and Control, 1998. Technical Report CVC TR-98-003, AIPS-98 Planning Competition Committee.
- [11] R. E. Fikes, N. J. Nilsson, Strips: A new approach to the application of theorem proving to problem solving, Artificial Intelligence 2 (1971) 189–208. doi:10.1016/0004-3702(71)90010-5.
- [12] D. Beßler, M. Pomarlan, M. Beetz, Owl-enabled assembly planning for robotic agents, in: Proceedings of the 2018 International Conference on Autonomous Agents (AAMAS '18), AAMAS '18, Stockholm, Sweden, 2018. Finalist for the Best Robotics Paper Award.
- [13] Q. Du, B. Li, Y. Du, S. Su, T. Fu, Z. Zhan, Z. Zhao, C. Wang, Fast task planning with neuro-symbolic relaxation, arXiv preprint arXiv:2507.15975 (2025).