An Extended Qualitative Model for Reasoning About 3D Objects Using Depth and Different Perspectives

Martí Ejarque
Computing Science Department, Umeå Universitet, Sweden

Zoe Falomir
Computing Science Department, Umeå Universitet, Sweden

Abstract

Reconstructing a three-dimensional structure from partial two-dimensional views requires spatial reasoning. This paper presents the EQ3D, a dimension-agnostic extension of the Qualitative 3D (Q3D) model that generalises qualitative volumetric reasoning to arbitrary $H \times W \times L$ configurations. EQ3D preserves symbolic depth representations while incorporating adjacent-view propagation rules that increase cross-view inference opportunities. Implemented in SWI-Prolog as a fully declarative pipeline, EQ3D infers hidden views from three orthographic `Front`–`Right`–`Up` views, producing correct inferred cells for the Q3D-Dataset which includes 24 objects of varying size and shape, while ambiguous cells remain explicitly undetermined. Coverage of hidden structure ranges from 26.85% to 82.67%, demonstrating that while inference remains logically sound, the proportion of determinable volume is inherently configuration-dependent. EQ3D provides an interpretable symbolic framework for qualitative 3D reasoning that scales across resolutions without size-specific adaptation.

2012 ACM Subject Classification Theory of computation → Automated reasoning; Theory of computation → Programming logic; Computing methodologies → Knowledge representation and reasoning; Computing methodologies → Spatial and physical reasoning

Keywords and phrases qualitative spatial reasoning, spatial cognition, first order logics, three-dimensional information, perspectives, object perception, SWI-Prolog

Digital Object Identifier 10.4230/LIPIcs.COSIT.2026.5

Supplementary Material
Software (Source Code): `https://github.com/Ejarque24/Extended-Q3D` [8]

Funding We acknowledge the funding by the Wallenberg AI, Autonomous Systems and Software Program (WASP) awarded by the Knut and Alice Wallenberg Foundation, Sweden.

1 Introduction

Qualitative Spatial and Temporal Reasoning (QSTR) models properties of space (e.g. topology, orientation, geometry) and their evolution through time. It produces representations and inferences over spatial and temporal configurations at a symbolic level [25, 2, 15]. Rather than relying on precise numerical coordinates, qualitative models capture structural properties such as direction, topology, adjacency, or relative depth [32]. This abstraction allows reasoning under incomplete or imprecise information, while remaining close to human understanding.

The qualitative description of three-dimensional objects and the inference of their spatial structure from partial views is a spatial problem. In many contexts, three-dimensional objects are described through a limited set of canonical projections [30]. Technical drawings and some educational spatial tasks present three orthogonal views (typically front, lateral and top) to represent an object [34] and ask to infer the occluded views. The inferential strength of some descriptions varies significantly: some sets of projections might constrain the hidden

structure of the object, while others can leave substantial ambiguity [23]. Understanding what makes a qualitative 3D description informative, and how much of the hidden structure can be reconstructed from it, is our research question.

The Qualitative 3D depth model (Q3D [10, 37]) addressed this challenge by introducing a symbolic representation of depth levels across three canonical views. In Q3D, each projection is described as a matrix of qualitative depth symbols, and logical consistency constraints govern the relations between opposing and adjacent views. The model demonstrated that hidden views can be partially reconstructed through symbolic reasoning alone, without recourse to metric geometry. In its original formulation, Q3D was defined over a fixed $3 \times 3 \times 3$ qualitative partition. This discretisation provides a conceptually transparent substrate for expressing depth relations and inter-view constraints. However, if qualitative depth reasoning is to apply to objects of varying size, the underlying constraints must be formulated independently of a specific dimension instantiation. Achieving such independence requires a parametrised reformulation in which depth semantics and cross-view consistency are invariant under arbitrary height (H), width (W), and length (L) dimensions, $H \times W \times L$.

Thus, this paper presents the Extended Qualitative (EQ3D), a dimension-agnostic reformulation of the Q3D framework, which generalises the qualitative volumetric description from a fixed $3 \times 3 \times 3$ partition to arbitrary $H \times W \times L$ dimensions, while preserving the ordinal interpretation of depth symbols and the logical structure of inter-view constraints. The extension is not achieved by merely scaling the original construction, but by redefining adjacency relations and boundary correspondences in an index-independent manner. As a result, the qualitative semantics of depth and consistency become invariant under changes in resolution and aspect ratio. By establishing this dimensional independence, EQ3D provides a unified qualitative framework capable of describing objects of arbitrary size within a single formal system. In addition, the generalised formulation makes the propagation of constraints across views explicit, enabling a systematic analysis of how structural configurations influence the reconstruction of hidden perspectives.

The rest of this paper is organised as follows. Section 2 describes related work. Section 3 outlines the Q3D model. Section 4 describes the Extended Q3D model and its inference mechanisms. Section 5 report the experimental results and coverage analysis. Section 6 discusses the results and structural factors influencing inferential reach. Finally, Section 7 concludes the paper and outlines directions for future work.

2 Related Work

Qualitative models [14] have been developed to solve cognitive spatial reasoning tests, such as object sketch recognition tests [27], oddity tasks [28], Raven's Progressive Matrices [29], paper folding reasoning tests [9, 13], qualitative shape composition [12, 31] and object rotation tests [11]. As far as we are concerned, there is no qualitative model developed to reason about the volumetric description of a generalized 3D object of $H \times W \times L$. In the field of Qualitative Spatial and Temporal Reasoning (QSTR), the most relevant calculi was reported by Dylla et al. [7], and there are no calculi that reasons about volumes, that we are aware of.

In the field of computer graphics and geometrical modeling, reconstructing three-dimensional structures/objects from multiple orthographic views has been studied extensively [16, 19]. Early symbolic and geometric approaches assume precise correspondences between views and rely on deterministic constraints to recover a unique 3D model [35]. Line-drawing interpretation and projection-consistency methods, for instance, compute spatial structure through exact geometric reasoning over edges, views, and junctions [21, 26, 38]. While effective under idealised conditions, such approaches presuppose metric precision and are not designed for qualitative abstraction or symbolic depth representations.

Some non-probabilistic "soft" approaches have also explored rule-based treatments of ambiguity while reconstructing 3D objects. Gorgani et al. [17] proposed a fuzzy surface analysis framework that, like EQ3D, relies on expert-defined rules rather than purely numeric optimisation. However, whereas their model represents uncertainty through graded membership within a fuzzy setting, EQ3D adopts a fully qualitative symbolic semantics in which ambiguity arises from logical underdetermination.

Recent work has reframed orthographic reconstruction as a data-driven inference problem. SPARE3D [18] introduces a benchmark for spatial reasoning over three-view drawings, including tasks such as view consistency checking, pose recognition, and shape generation. While these tasks are closely related to EQ3D in that it also starts from `Front-Right-Up` views, SPARE3D relies on learned statistical patterns rather than explicit symbolic constraints. Notably, standard deep neural models perform close to random guessing on several reasoning tasks [18], highlighting the difficulty of capturing structural spatial constraints implicitly. Because SPARE3D does not provide qualitative depth encodings, it cannot be directly used within the EQ3D reasoning framework.

PlankAssembly [20] employs transformer-based architectures to predict "shape programs" for block-like CAD objects from orthographic views, similarly to EQ3D. By learning structural priors and encoding constraints implicitly in the decoding process, the system achieves robustness to noise and incomplete inputs. However, unlike EQ3D's explicit symbolic rules and declarative queryable inference mechanisms, its consistency conditions and inference rules remain embedded in model parameters and not directly explainable or inspectable.

Hybrid neuro-symbolic approaches have also been explored in block-based reasoning domains [24, 4]. Krishnaswamy, Friedman, and Pustejovsky [24] combined qualitative spatial relations with learning and heuristic search to acquire structural concepts from sparse and noisy examples. However, there is not a neuro-symbolic model for constructing 3D objects from 2D views, as far as we are concerned.

3 Preliminaries: the Qualitative 3D depth model

When reasoning about real objects in space, humans naturally think in terms of three-dimensional volumes and multiple perspectives. The Q3D model [10] formalises this intuition by encoding orthographic projections through symbolic depth values rather than numeric measurements. The Q3D depth Descriptor [10, 37] of an object is defined by three views:

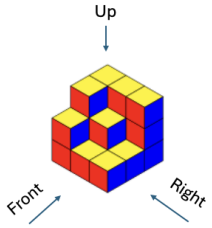
$$\begin{aligned} &view(F, Object, H, W, Q3Dmatrix(H, W)),\\ &view(R, Object, H, L, Q3Dmatrix(H, L)),\\ &view(U, Object, L, W, Q3Dmatrix(L, W)). \end{aligned}$$

where F, R, and U denote the `Front`, `Right`, and `Up` perspectives, while H, W, and L represent the height (H), width (W), and length (L) of the object's bounding box. Each matrix entry belongs to the symbolic set:

$$Q3Dmatrix(D_1, D_2) \subseteq \{a, b, c, d, \ldots, *\}.$$

where D_1 and D_2 are the dimensions of the matrix (i.e. pairs of 3D dimensions: H, W and H, L and L, W) and the elements of each matrix describe ordinal depth layers composed of equally sized volumetric units: a denotes the surface of the object, b the first level of depth (one unit removed), c the second level of depth (two units of volume removed), d the third level of depth (three units of volume removed) and so on. The symbol $*$ indicates that all volumes along a depth column have been removed, so there is a hole. These symbols encode relative ordering rather than metric distance.

Figure 1 presents an object described by the Q3D where volumes of a specific size produced a $3 \times 3 \times 3$ grid with volumetric features for the `Front`, `Right`, and `Up` sides. Note that another volumetric size could have been used producing other dimensions for the object descriptor (e.g. $6 \times 6 \times 6$).



Q3D descriptor for this 3x3x3 object:

`Front`	`Right`	`Up`
$[b, c, c]$	$[*, c, a]$	$[a, a, a]$
$[a, b, c]$	$[c, b, a]$	$[a, b, c]$
$[a, a, a]$	$[a, a, a]$	$[b, c, c]$

Logically as Prolog facts:

$view(front, obj, 3, 3, [[b, c, c], [a, b, c], [a, a, a]]).$
$view(right, obj, 3, 3, [[*, c, a], [c, b, a], [a, a, a]]).$
$view(up, obj, 3, 3, [[a, a, a], [a, b, c], [b, c, c]]).$

Figure 1 Example of (a) an object observed from `Front`, `Right`, and `Up` sides; (b) an intuitive representation of the Q3D descriptor; (c) the logic facts produced by the Q3D descriptor for this specific object.

Consistency Checking across Views

Beyond representation, Q3D imposes geometric compatibility constraints between views. A Q3D description is *consistent* [10] if all pairwise projections can be jointly realised as orthographic views of a single three-dimensional object:

$$\begin{aligned} consistent\_Q3D(F, R, U) \leftrightarrow\ & consistent\_perspective(F, R) \wedge \\ & consistent\_perspective(F, U) \wedge \\ & consistent\_perspective(R, U). \end{aligned}$$

Intuitively, two views are consistent if there exists (at least) a 3D arrangement of volumetric blocks whose projections simultaneously satisfy the qualitative depth relations encoded in both views. Thus, `Front`–`Right` consistency requires that the volume values in last column of the `Front` view are consistent with the depth columns of the `Right` view, since they meet in an object edge and therefore these volumes are the same seen from both views. `Front`–`Right` consistency implies `Right`–`Front` consistency, which is defined as:

$$\begin{aligned} F_{i,3}(a) &\leftrightarrow R_{i,1}(a) \\ F_{i,3}(b) &\leftrightarrow R_{i,2}(a) \wedge (R_{i,1}(b) \vee R_{i,1}(c) \vee R_{i,1}(*)) \\ F_{i,3}(c) &\leftrightarrow R_{i,3}(a) \wedge (R_{i,1}(b) \vee R_{i,1}(c) \vee R_{i,1}(*)) \wedge (R_{i,2}(b) \vee R_{i,2}(c) \vee R_{i,2}(*)) \\ F_{i,3}(*) &\leftrightarrow \bigwedge_{k=1}^{3} \neg R_{i,k}(a). \end{aligned}$$

where $\forall k, i \in \{1, 2, 3\}$. Similarly, `Front`–`Up` consistency and `Right`–`Up` consistency are defined.

Inference of Depth Values in Occluded Object Sides

The Q3D model also provided deterministic rules for inferring depth values in the occluded sides of an object: `Back`, `Left`, and `Down`. For that, these rules exploit space continuity and depth interrelations. For example, `Front` and `Back` are opposite sites, if the Q3D include an item indicating the last level of depth in a view (e.g., level c in `Front` in an object of $3 \times 3 \times 3$ dimension), that involves the existence of the first level of depth (a) in the opposite view,

that is `Back`. The same interrelations appear in opposite `Right` and `Left` sides and in `Up` and `Down`. Regarding space continuity, the case of holes involves that symbols ($*$) propagate from-to opposite sides of an object.

In summary, the Q3D depth model introduced: (i) a symbolic volumetric representation, (ii) cross-view consistency constraints, and (iii) inference mechanisms based on space continuity (e.g. hole continuity) and interrelations (e.g. depth relativity). It successfully showed that spatial structure and depth relations can be inferred from complementing perspectives. The Q3D reasoning model was implemented and tested in a SWI-Prolog application using 3×3×3 objects [10] and then it was implemented in a videogame [37] with the goal of training players' spatial reasoning skills.

In this paper, we extend the Q3D model to any dimension and we also include a new inference rule for discovering volumes in occluded sides which is based on conceptual neighbourhood between sides.

4 The Extended Qualitative 3D Depth Model (EQ3D)

The Extended Qualitative 3D depth model (EQ3D) builds directly upon the foundations established by the original Q3D model [10] by providing descriptors for objects of any dimension $H \times W \times L$, preserving the qualitative initial representation while extending its expressive power to a broader range of object geometries. This generalization enables finer volumetric partitions and supports shapes that are elongated, asymmetric, or otherwise more detailed than those captured by the original configuration.

Specifically, EQ3D generalizes to arbitrary object size ($H \times W \times L$) while preserving the underlying qualitative depth semantics, it propagates depth information using adjacency-based cross-view constraints across neighbouring views.

Moreover, it exploits adjacency relations between views to further extend the inference capabilities of Q3D by using the consistency checking rules as inference rules discovering volumes existing in the occluded neighbouring sides (e.g. for inferring new depth values for `Back`, we can check the views `Right` vs. `Back`, `Left` vs. `Back`, `Down` vs. `Back` and `Up` vs. `Back`, specifically at the edges of the object, where the volumes seen are the same.) These new rules exploit local depth continuity across shared boundaries, allowing information to propagate between neighbouring views. In doing so, EQ3D complements and enriches the original inference framework, extending qualitative reasoning towards more complex and realistic 3D interpretations. These generalizations provide a unified qualitative framework that supports both consistency checking and the reconstruction of hidden views from partial input.

Notation and Preliminaries. EQ3D retains the original formal foundations of Q3D, where three canonical perspectives (`Front` (F), `Right` (R), and `Up` (U)) encode qualitative depth using symbolic letters. Each perspective is expressed as a matrix whose entries belong to: $\Sigma = \Sigma_d \cup \{*\}$, $\Sigma_d = \{a, b, c, d, \dots\}$ where a denotes the nearest visible surface, deeper levels follow the alphabetical order, and $*$ represents enforced absence of volume.

The qualitative structure is now generalized to an arbitrary grid of dimensions (H, W, L), corresponding respectively to the vertical, horizontal, and depth axes. Each axis is discretized into uniform volumetric units, without fixing the number a priori, allowing the model to adjust to the complexity of the described object.

Formally, a perspective P is a matrix of depth symbols: $P = [p_{i,j}]$, $p_{i,j} \in \Sigma$, with dimensions determined by how the view aligns with the (H, W, L) grid.

▶ **Definition 1.** *Let the qualitative reference system of EQ3D be:*

$$EQ3D^{(H,W,L)} = \left\{P_v \mid v \in V,\ P_v \subseteq \Sigma^{n_v \times m_v},\ (n_v, m_v) \in \{(H,W),(H,L),(L,W)\}\right\},$$

representing all canonical perspectives with their geometrically determined sizes.

The depth alphabet Σ_d encodes only ordinal information (near/far ordering), independent of the chosen grid resolution. This separation ensures that qualitative reasoning is scale-invariant across different object sizes.

Generalization to Arbitrary Dimensions $H \times W \times L$. The EQ3D generalizes dimensions by expressing all adjacency relations between views in an index-agnostic manner: the rules refer only to the boundary indices that two views share, and therefore remain valid independently of the absolute size or dimension.

In the EQ3D qualitative partition, each canonical view corresponds to a projection onto one of the coordinate planes. Adjacent views meet along a one-dimensional boundary, and the structure of this boundary determines how their symbolic descriptions must agree. For example, the rightmost column of the `Front` view coincides with the leftmost column of the `Right` view; similarly, the top row of the `Front` view aligns with the bottom edge of the `Up` view; and the upper row of the `Right` view aligns with the rightmost edge of the `Up` view, subject to the natural mirroring imposed by the viewing direction.

▶ **Definition 2** (Neighbouring boundary alignments)**.** *Let $F \in N_{\text{depths}}^{H\times W}$ and $R \in N_{\text{depths}}^{H\times L}$ denote the* ***Front*** *and* ***Right*** *views. For each row i, the boundary cells $F(i,W)$ and $R(i,1)$ describe the same vertical line of the 3D object. If $F(i,W) = *$, then the entire $R(i,1)$ row must be free of volume units, since there is a hole on the* ***Front*** *view at that location which excludes the possibility of a visible surface from the* ***Right*** *view. Conversely, if $F(i,W) \neq *$, then row i of the* ***Right*** *view must show a surface at exactly the depth index associated with that qualitative depth value. A symmetric constraint maps the leftmost column of R back to the appropriate mirrored column of F, ensuring that both views encode the same shared edge consistently. The remaining view adjacency pairs in the object follow the same principle, thus:*

$$\begin{aligned} &F(i,W) \leftrightarrow R(i,1), && i = 1,\dots,H,\\ &F(1,j) \leftrightarrow U(1,j), && j = 1,\dots,W,\\ &R(1,k) \leftrightarrow U(k,W), && k = 1,\dots,L. \end{aligned}$$

That is, the top row of the `Front` view shares its boundary with the nearest edge of the `Up` view; thus, each symbol $F(1,j)$ constrains the corresponding column of U when read from bottom to top. Likewise, the upper row of the `Right` view meets the rightmost column of the `Up` view; here, the alignment requires horizontal mirroring due to the opposite viewing direction, but the constraint itself is structurally identical to the previous cases. Since the rules quantify only over these boundary indices, and not over fixed-size patterns, then the adjacency constraints automatically generalize to any choice of (H,W,L). Thus, EQ3D preserves the structure of Q3D consistency reasoning while allowing objects of arbitrary resolution and aspect ratio to be described and validated.

Computational implications. Checking pairwise edge consistency across the three adjacent pairs ($F \leftrightarrow R$, $F \leftrightarrow U$, $R \leftrightarrow U$) requires a number of comparisons proportional to the lengths of shared edges. Summed over all rows/columns, this yields time linear in the total number of view cells, $\mathcal{O}(H \cdot W + H \cdot L + L \cdot W)$. When inference propagates information throughout

the interior of the volume (e.g., to reconstruct occluded cells or to run completion on all views), the work scales linearly with the number of volumetric units, $\mathcal{O}(H \cdot W \cdot L)$. Therefore, EQ3D preserves linear scaling with problem size at both the surface (views) and volume levels.

Illustrative example (4×4×3). Consider an object discretized over an EQ3D grid with dimensions $H = 4$ (Height, Cartesian axis y), $W = 4$ (Width, Cartesian axis x), and $L = 3$ (Length, Cartesian axis z). Figure 1 shows the qualitative description of such an object using its three canonical perspectives: `Front` (4×4), `Right` (4×3), and `Up` (3×4).

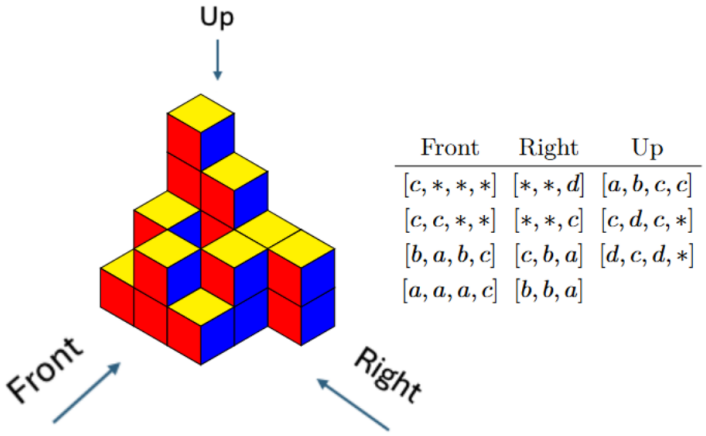


Figure 2 Example of a discretized 3D object showing the `Front`, `Right`, and `Up` views.

Figure 2 shows the EQ3D descriptor for an object of size: 4×4×3. The `Front` view specifies, for each of the 4 rows and 4 columns, the nearest visible depth letter along the front-to-back axis (i.e. Cartesian z axis). The `Right` view provides the same information from the right-hand side: for each of the 4 rows and 3 columns, it encodes the nearest visible depth along the right-to-left axis (i.e. Cartesian x axis). The `Up` view, of size 3×4, corresponds to looking from above (i.e. Cartesian y axis): its 3 rows index depth layers from nearest to farthest, while its 4 columns align with the horizontal positions shared with the Front view.

4.1 EQ3D Consistency Checking

The EQ3D model checks that the `Front`, `Right` and `Up` descriptors are logically coherent between them to produce a robust description of a single 3D object. This is done by enforcing three layers of consistency constraints: well-formed views, object completeness and cross-view consistency along shared edges.

A view is well-formed if all rows have the same length, its declared dimensions (*rows*, *cols*) match the actual matrix dimensions, and every entry belongs to N_{depths}. Views that violate any of these conditions are rejected as invalid descriptions and cannot be used for EQ3D reasoning. Completeness is considered when all size canonical perspectives are available. In practice, our focus is on the three input views F, R, U, while B, L, D is used as ground truth for the later inferred views. However, whenever all six views are provided completeness checking ensures that no view is missing and no extra/non-canonical label is present. This guarantees that all consistency rules below are defined on a fixed, common set of perspectives.

Adjacency-based consistency. The core of EQ3D consistency checking is the set of *adjacency rules* that relate pairs of neighbouring views along their shared edges. Conceptually, these rules state that whenever two views touch in 3D, their boundary cells must encode compatible depth information. EQ3D checks three adjacency pairs, each in both directions: (F, R), (F, U), (R, U) yielding the six constraints FR, RF, FU, UF, RU, and UR. In all rules below, we assume that each depth letter $\ell \in \{a, b, c, \dots\}$ is mapped to a *depth index* $\mathrm{idx}(\ell) \in \mathbb{N}$, starting at 0 for a, 1 for b, and so on.

▶ **Definition 3** (EQ3D Consistency Front → Right)**.** *For each row i, the last column of the Front view (rightmost edge) meets the first column of the Right view (leftmost edge). The rule is:*

(FR-)* $F(i, W) = * \;\Rightarrow$ *no cell* $R(i, \cdot)$ *in that row is* a,
(FR-a) $F(i, W) = \ell \neq * \;\Rightarrow\; R\big(i, \mathrm{idx}(\ell)\big) = a.$

Intuitively, if the last visible point in `Front` is a hole, `Right` must not see a surface a along that row. If instead Front sees depth ℓ, then Right must show a surface at the corresponding depth index.

▶ **Definition 4** (EQ3D Consistency Front ← Right)**.** *Symmetrically, we check the first column of Right against the entire row of Front, but now taking into account horizontal mirroring. Let W be the number of columns of Front, and* $\mathrm{mir}(k) = W - 1 - k$ *the mirrored index. Then let* $\mathrm{mir}(k) = W - 1 - k$ *denote the horizontal mirroring index for a view with W columns. For each row i:*

(RF-)* $R(i, 1) = * \;\Rightarrow$ *no cell* $F(i, \cdot)$ *in that row is* a,
(RF-a) $R(i, 1) = \ell \neq * \;\rightarrow\; F\big(i, \mathrm{mir}(\mathrm{idx}(\ell))\big) = a.$

FR and RF together ensure that Front and Right encode the same edge in a consistent way, up to the natural mirroring of the viewing axes.

▶ **Definition 5** (EQ3D Consistency Front → Up)**.** *The top row of Front shares its boundary with the front edge of the Up view. For each column j, we compare $F(1, j)$ with the column of Up at position j when read from bottom to top:*

(FU-)* $F(1, j) = * \;\Rightarrow$ *no cell in the corresponding Up column is* a,
(FU-a) $F(1, j) = \ell \neq * \;\Rightarrow$ *the Up column at j has a at depth index* $\mathrm{idx}(\ell)$.

Thus, a hole at the top of `Front` discards surfaces directly above it in `Up`, whereas a visible depth letter demands a matching surface at the corresponding vertical dimension Length (Cartesian axis z) in the Up view.

▶ **Definition 6** (EQ3D Consistency Front ← Up)**.** *Conversely, the bottom row of Up is aligned with the full height of Front. For each column j:*

(UF-)* $U(L, j) = * \;\Rightarrow$ *no cell in the Front column j is* a,
(UF-a) $U(L, j) = \ell \neq * \;\rightarrow\; F\big(\mathrm{idx}(\ell) + 1,\, j\big) = a.$

Here, a hole at the bottom of Up excludes any visible surface in the corresponding Front column, while a depth letter fixes the row at which Front must show a surface.

▶ **Definition 7** (EQ3D Consistency Right → Up)**.** *The top row of Right is adjacent to the right edge of Up. For each depth index that is simultaneously valid for both views, the rule is:*

(RU-)* $R(1, k) = * \;\Rightarrow$ *the corresponding mirrored row of Up has no* a,
(RU-a) $R(1, k) = \ell \neq * \;\Rightarrow$ *the corresponding mirrored row of Up contains a at index* $\mathrm{idx}(\ell)$.

Each position on top edge of Right must be compatible with the row of Up that represents the same depth, read with the appropriate horizontal mirroring.

▶ **Definition 8** (EQ3D Consistency Right ← Up)**.** *The rightmost column of Up is matched with a family of depth-indexed columns in Right. For each row i of Up (which corresponds to a depth layer), we enforce:*

$$\text{(UR-*)} \quad U(i, W_U) = * \;\Rightarrow\; \text{no } a \text{ occurs in the corresponding depth column of Right},$$
$$\text{(UR-}a\text{)} \quad U(i, W_U) = \ell \neq * \;\Rightarrow\; \text{the Right column at depth } \mathrm{idx}(\ell) \text{ contains } a \text{ at row } i.$$

An EQ3D object is said to be adjacency-consistent if all rules (Front ↔ Right ∧ Front ↔ Up ∧ Right ↔ Up) are satisfied simultaneously. In logical terms, each rule is a local constraint over pairs of boundary cells, but their conjunctions ensure that the three given views F, R, U could arise from a single coherent 3D configuration. This global consistency is a prerequisite for the following inference stage, which uses, to some extent, the same qualitative set and adjacency relations to complete the hidden views B, L, D.

4.2 EQ3D Reasoning Inferences

Assuming the orthographic views `Front`, `Right` and `Up` are provided in an EQ3D description, we ask: how much of the remaining object can be deduced logically? In particular, can the `Back`, `Left` and `Down` perspectives be partially computed from the given views?

Each view is represented by a matrix of symbols over $\{a, b, \ldots\} \cup \{*\}$, where letters encode depth along a sightline (a nearest to the observer) and `*` denotes a hole (an enforced absence). Unknown or not yet inferred entries are written as "_":

EQ3D reasons about hidden views using three families of qualitative inference rules: opposite-view depth relativity, opposite-view hole continuity, and additional adjacent depth cues that directly constrain `Back`, `Left` and `Down`.

4.2.1 Opposite-view depth relativity

Depth-relativity captures the idea that the deepest visible symbol along one axis guarantees the presence of a surface in the opposite view. EQ3D distinguishes three depth axes:

- deep_{FB}: deepest letter along the `Front`/`Back` axis (i.e. Cartesian z axis).
- deep_{LR}: deepest letter along the `Left`/`Right` axis (i.e. Cartesian x axis).
- deep_{UD}: deepest letter along the `Up`/`Down` axis (i.e. Cartesian y axis).

Intuitively, depth relativity propagates guaranteed occupancy, hole continuity propagates guaranteed absence, and adjacent depth cues exploit shared boundaries to refine the structure of hidden views. For a given object, the maximum number of depth levels are determined based on the dimension of the axis (e.g. if there are four levels, $\mathrm{deep}_{LR} = d$).

▶ **Definition 9** (Inference Depth Relativity Front ↔ Back)**.** *Let F be the Front view with N columns, and B the Back view. For each row i and column j:*

$$F(i, j) = \mathrm{deep}_{LR} \;\leftrightarrow\; B\big(i, \mathrm{mir}(j)\big) = a, \quad \mathrm{mir}(j) = H - 1 - j.$$

That is, whenever Front sees the deepest possible depth at position (i, j), the mirrored column $\mathrm{mir}(j)$ in the Back row i must contain a visible surface a.

▶ **Definition 10** (Inference Depth Relativity Right ↔ Left). *Symmetrically, let R be the Right view with N columns, and L the Left view. For each row i and column j:*

$$R(i,j) = \text{deep}_{LR} \;\leftrightarrow\; L\big(i, \text{mir}(j)\big) = a.$$

Again, a deepest letter on Right at (i,j) enforces a visible surface on Left at the mirrored position of the same row.

▶ **Definition 11** (Inference Depth Relativity Up ↔ Down). *For the vertical axis, let U be the Up view with K rows and M columns, and D the Down view. For each Up cell (i,j):*

$$U(i,j) = \text{deep}_{UD} \;\leftrightarrow\; D\big(\text{mir}_v(i), j\big) = a, \quad \text{mir}_v(i) = L - 1 - i.$$

Here, a deepest letter seen from above at row i, column j implies a surface at the mirrored vertical depth $\text{mir}_v(i)$ in the Down view at the same horizontal position j. Intuitively, if the object is "full" as far as the Up view can see at that sightline, then the opposite Down view must also encounter a surface at the extreme depth.

These depth-relativity rules never introduce holes; they only add certain a cells on the opposite view, reflecting the idea that reaching the last usable depth from one side guarantees occupancy when seen from the other side.

4.2.2 Opposite-view hole continuity

Hole continuity states that a hole observed from one side of the object must also appear as a hole in the directly opposite view, at the mirrored coordinate. This captures the idea that an enforced absence of volume cannot be "filled" when seen from the other side. For each axis, the deepest letter is determined by the maximum number of qualitative depth levels implicit in the corresponding view dimensions. For instance, if four levels are used along the left–right axis, then $\text{deep}_{LR} = d$.

▶ **Definition 12** (Inference Hole Propagation Front ↔ Back). *Let F have H columns. For every cell (i,j):*

$$F(i,j) = * \;\leftrightarrow\; B\big(i, \text{mir}(j)\big) = *.$$

▶ **Definition 13** (Inference Hole Propagation Right ↔ Left). *Similarly, for Right and Left:*

$$R(i,j) = * \;\leftrightarrow\; L\big(i, \text{mir}(j)\big) = *.$$

▶ **Definition 14** (Inference Hole Propagation Up ↔ Down). *For Up and Down, using vertical mirroring:*

$$U(i,j) = * \;\leftrightarrow\; D\big(\text{mir}_v(i), j\big) = *.$$

Opposite view hole continuity and depth-relativity are complementary: the former propagates enforced absences ($*$), while the latter propagates guaranteed surfaces (a) from the deepest observable level. Both act monotonically: they never remove information, only refine unknown cells "_" into either $*$ or a.

4.2.3 Additional adjacent depth cues

Beyond the opposite-view constraints, EQ3D incorporates a set of adjacent depth cues that arise from views sharing not only edges but also orthographic sightlines whose extremal rows or columns encode depth information. These cues exploit the fact that certain boundary positions (such as the last column of the Right view, the top row of the `Up` view, or the leftmost column of `Up`) correspond directly to depth indices in the hidden views. As a result, specific depth letters in these boundary positions can be interpreted as direct constraints on the `Back`, `Left`, or `Down` views.

Concretely, a depth letter appearing in an extremal row or column of a visible view fixes the location of a surface cell in the corresponding hidden view. For instance, the last column of the **Right** view represents the furthest depth positions along the right-to-left axis; thus, if $R(i, L) = \ell \neq *$, the depth index $\text{idx}(\ell)$ specifies the column of the `Back` view in which a surface must appear at row i: $R(i, L) = \ell \neq * \;\rightarrow\; B\big(i, \text{idx}(\ell)\big) = a$.

Analogous forms of depth-index projection apply to the top row of the `Up` view (informing `Back`), the leftmost column of `Up` (informing `Left`), and the bottom rows of `Right` or `Front` (informing `Down`). These cues do not replace the adjacency rules but enrich them by providing direct depth-to-position mappings that become available whenever the relevant boundary cells contain explicit depth letters. Together with the previous inference mechanisms, they contribute to a more complete and deterministic reconstruction of the hidden views B, L, and D.

5 Results

Dataset and Experimental Setup

The EQ3D model was implemented as first-order Horn clauses using SWI-Prolog[1] as the testing platform [36] following a fully declarative inference pipeline. All experiments were executed under identical symbolic parameters, without size-dependent hardcoding, ensuring generality across $H \times W \times L$ configurations.

Then it was evaluated on the Q3D-Dataset, which contains 24 objects ranging from compact cubic configurations ($2 \times 2 \times 2$) to larger rectangular shapes (up to $6 \times 7 \times 6$) which was originally created for the Q3D-videogame [37] and it is a compilation of objects from the literature [1, 3, 34, 10]. Each object was specified through six canonical perspectives (`Front`, `Right`, `Up`, `Back`, `Left`, `Down`), following the EQ3D reference system introduced in Section 4.

For each object, only the three visible views (`Front`, `Right`, `Up`) were provided as input facts. The remaining three hidden views (`Back`, `Left`, `Down`) were inferred automatically by the EQ3D model using the adjacency and depth-consistency rules defined in Section 4.1. Objects were normalised prior to inference to ensure canonical alignment and to avoid ambiguities due to symmetry or rotation. All experiments were executed under identical symbolic parameters, without size-dependent hardcoding, ensuring generality across $H \times W \times L$ configurations. Both the data and the code for this work are available here.[2]

[1] SWI-Prolog: `http://www.swi-prolog.org/`
[2] `https://github.com/Ejarque24/Extended-Q3D.git`

Evaluation Protocol

We evaluated the EQ3D model using two measures: *correctness* and *inferential completeness*. Correctness is assessed by comparing every inferred hidden-view cell against its ground-truth qualitative depth value for each object. Inferential completeness is assessed using a coverage measure which is defined over the hidden views only (`Back`, `Left`, `Down`), and reported alongside compact inference indicators.

Correctness. Across all objects, inferred cells match the ground-truth qualitative descriptors with **100% accuracy**. This follows from the soundness of the EQ3D inference rules: the system derives only information that logically follows from the visible view together with volumetric continuity and depth relativity. No contradictions or inconsistent configurations were produced.

Coverage as Inferential Completeness. To quantify inferential reach, we measure how much of the hidden structure (`Back`, `Left`, `Down`) can be determined from the visible views. We defined coverage as the proportion of hidden-view cells whose values become determined (either a depth symbol *a,b,c,...* or a hole `*`). The hidden-view cells are composed of all the unseen cells in `Back`, `Left`, `Down`, including the boundary cells shared with the shown views:

$$\text{Coverage}(O) = \frac{\texttt{determined hidden cells}(O)}{\texttt{total hidden cells}(O)} \tag{1}$$

Thus, coverage quantifies *inferential completeness*, independently of correctness. Coverage values range from **26.85%** (object 18) to **82.67%** (object 9), with a dataset mean of approximately 60%. This confirms that while inference is uniformly correct, the proportion of reconstructable hidden structure varies substantially across configurations.

To contextualise coverage, we report three compact global indicators that capture inferential behaviour and visible structural properties. Table 1 shows representative objects spanning the coverage spectrum. The reported indicators are defined as follows:

- **Collision Rate (CR):** inferential redundancy, measured as the proportion of inference events that do not expand the set of determined hidden cells (i.e., multiple rule firings targeting already-determined cells). Lower values indicate broader spatial dispersion of inference over distinct target cells.
- **Boundary Share (BSS):** the proportion of inference events triggered by the dedicated inference-boundary rules (Right→Back last-column, Up→Back top-row, Up→Left left-column, Right→Down last-row, Front→Down last-row). High BSS indicates that inference is dominated by boundary-propagation rules rather than interior propagation.
- **INER (Interior Non-Extreme Ratio):** the proportion of visible cells *excluding inference-boundary cells* whose values are neither holes (`*`) nor view-dependent maximum depth symbols. INER provides a compact measure of how much interior visible structure lacks strong extremal cues.

Coverage Spectrum

Objects cluster into three empirical types:

- **High coverage** ($\geq 70\%$; 7 objects): objects 5, 9, 10, 11, 12, 13, and 21, with object 9 reaching 82.67%.
- **Medium coverage** (50–70%; 9 objects): objects 1, 2, 3, 4, 6, 7, 8, 14, 16, 17, and 24.
- **Low coverage** ($< 40\%$; 4 objects): objects 18, 20, 22, and 23.

Table 1 Coverage and selected inference indicators.

Object	Origin	Coverage (%)	Collision Rate (CR)	Boundary Share (BSS)	INER
o1	Q3D [10]	55.56	0.3750	0.6250	0.4286
o2	Q3D [10]	59.26	0.3846	0.5769	0.2857
o3	Q3D [10]	55.56	0.1176	0.8824	0.8571
o4	Q3D [10]	51.85	0.1765	0.8824	0.9286
o5	DST_01 [34]	74.07	0.3548	0.4516	0.0714
o6	DST_02 [34]	55.56	0.2500	0.4500	0.6429
o7	DST_03 [34]	66.67	0.2800	0.3600	0.3571
o8	PAT_5A [3]	66.67	0.1351	0.2162	0.4333
o9	PAT_5B [3]	82.67	0.0746	0.1642	0.1923
o10	PAT_5C [3]	79.17	0.2830	0.3585	0.0333
o11	PAT_5D [3]	72.92	0.1463	0.2927	0.3000
o12	PAT_5Z [3]	75.00	0.5263	0.5263	0.0000
o13	PAT_100 [3]	76.00	0.1618	0.3529	0.3462
o14	PAT_113 [3]	58.33	0.2821	0.3590	0.3333
o15	PAT_222 [3]	48.15	0.1333	0.8667	1.0000
o16	PAT_224 [3]	61.33	0.0800	0.1800	0.4808
o17	PAT_225 [3]	68.75	0.1316	0.4211	0.4000
o18	PAT_230 [3]	26.85	0.0333	0.5333	0.9750
o19	PAT_219 [3]	49.33	0.0750	0.6000	0.7115
o20	PAT_114 [3]	38.27	0.0882	0.7647	0.9074
o21	PAT_200 [3]	79.17	0.0952	0.8095	0.5556
o22	PAT_220 [3]	38.55	0.0857	0.6286	0.8621
o23	PAT_221 [3]	38.89	0.0667	0.8000	0.9167
o24	PAT_689 [3]	52.78	0.0952	0.7143	0.8889

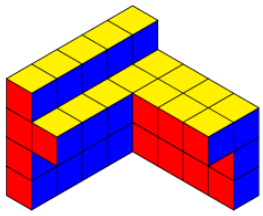
These results demonstrate that EQ3D inference generalises across dimensional configurations while preserving correctness. However, inferential completeness is inherently configuration-dependent: identical rule sets yield markedly different coverage levels depending on the distribution of visible depth information. This variability is not random but systematically linked to the balance between boundary-derived cues and interior visible structure. The mechanisms underlying this behavior are analysed in the following discussion.

Inference Output

All inferred matrices for the full set of twenty-four objects, including the six canonical views (FRU and BLD) and the cell-level inference outputs, are released as part of the open EQ3D dataset. For each object, the dataset contains: (i) the visible input views, (ii) the inferred hidden faces, (iii) the corresponding ground-truth matrices, and (iv) cell-level markers indicating determined and undetermined entries. A couple of examples can be seen in Figure 3.

This enables full reproducibility of all reported coverage values, independent verification of correctness, and detailed inspection of which cells are logically entailed by the visible projections and which remain underdetermined. The dataset makes explicit that undetermined cells are not inference errors, but geometric ambiguity under orthographic constraints.

o13



FRU

Front	Right	Up
$[*,*,*,*,*]$	$[*,*,*,*,*]$	$[b,c,c,c,c]$
$[a,*,*,*,*]$	$[e,e,e,e,e]$	$[b,c,c,c,c]$
$[a,a,d,d,d]$	$[d,d,d,a,a]$	$[b,c,*,*,*]$
$[a,e,e,e,e]$	$[e,e,e,e,a]$	$[b,c,*,*,*]$
$[a,e,e,e,e]$	$[e,e,e,e,a]$	$[b,c,*,*,*]$

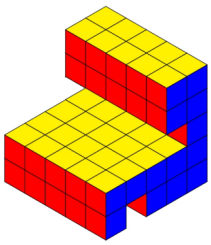
Inferred

Back	Left	Down
$[*,*,*,*,*]$	$[*,*,*,*,*]$	$[a,\_,*,*,*]$
$[*,*,*,*,a]$	$[a,a,a,a,a]$	$[a,\_,*,*,*]$
$[a,a,a,a,\_]$	$[\_,\_,\_,\_,\_]$	$[a,\_,*,*,*]$
$[a,a,a,a,\_]$	$[\_,a,a,a,a]$	$[a,\_,\_,\_,\_]$
$[a,a,a,a,\_]$	$[\_,a,a,a,a]$	$[\_,a,a,a,a]$

BLD (GT)

Back	Left	Down
$[*,*,*,*,*]$	$[*,*,*,*,*]$	$[a,c,*,*,*]$
$[*,*,*,*,a]$	$[a,a,a,a,a]$	$[a,c,*,*,*]$
$[a,a,a,a,a]$	$[a,a,a,a,a]$	$[a,c,*,*,*]$
$[a,a,a,a,a]$	$[a,a,a,a,a]$	$[a,c,c,c,c]$
$[a,a,a,a,a]$	$[a,a,a,a,a]$	$[a,a,a,a,a]$

o19



FRU

Front	Right	Up
$[d,d,d,d,d]$	$[*,*,*,a,a]$	$[a,a,a,a,a]$
$[d,d,d,d,d]$	$[*,*,*,a,a]$	$[a,a,a,a,a]$
$[e,e,e,e,e]$	$[*,*,*,*,a]$	$[d,d,d,d,d]$
$[a,a,a,a,a]$	$[a,a,a,a,a]$	$[d,d,d,d,d]$
$[a,a,a,a,a]$	$[a,*,a,a,a]$	$[d,d,d,d,d]$

Inferred

Back	Left	Down
$[a,a,a,a,a]$	$[a,a,*,*,*]$	$[a,a,a,a,a]$
$[a,\_,\_,\_,\_]$	$[\_,\_,*,*,*]$	$[\_,\_,\_,\_,\_]$
$[a,a,a,a,a]$	$[\_,*,*,*,*]$	$[\_,\_,\_,\_,a]$
$[a,\_,\_,\_,\_]$	$[\_,\_,a,a,a]$	$[\_,\_,\_,\_,a]$
$[a,\_,\_,\_,\_]$	$[\_,\_,\_,*,\_]$	$[\_,\_,\_,\_,a]$

BLD (GT)

Back	Left	Down
$[a,a,a,a,a]$	$[a,a,*,*,*]$	$[a,a,a,a,a]$
$[a,a,a,a,a]$	$[a,a,*,*,*]$	$[b,b,b,b,b]$
$[a,a,a,a,a]$	$[a,*,*,*,*]$	$[a,a,a,a,a]$
$[a,a,a,a,a]$	$[a,a,a,a,a]$	$[a,a,a,a,a]$
$[a,a,a,a,a]$	$[a,a,a,*,a]$	$[a,a,a,a,a]$

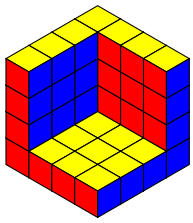
Figure 3 Inference comparison for objects o13 (76.00% coverage) and o19 (49.33% coverage).

6 Discussion

Scalability and Structural Generality

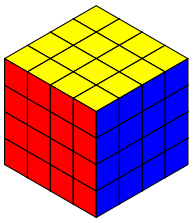
The experimental results confirm a central design claim of EQ3D: *logical soundness is invariant under dimensional scaling*. Across all tested configurations (from compact $2 \times 2 \times 2$ cubes to larger $6 \times 7 \times 6$ rectangular volumes) the inference rules required no modification, tuning, or size-specific adaptation. The rule system operates purely over symbolic adjacency relations, view correspondences, and ordinal depth constraints. As a consequence, scalability in EQ3D is *structural* rather than procedural: increasing grid resolution expands the domain of indices, but does not alter the reasoning mechanism itself.

At the same time, inferential *completeness* (measured through coverage) varies significantly across objects. While correctness remains constant (100% accuracy), coverage ranges from 26.85% (object 18) to 82.67% (object 9). This variability does not reflect instability; instead, it reveals that the proportion of determinable hidden structure is inherently dependent on qualitative configuration.



Front	Right	Up
$[a, d, d, d]$	$[d, d, d, a]$	$[a, a, a, a]$
$[a, d, d, d]$	$[d, d, d, a]$	$[a, d, d, d]$
$[a, d, d, d]$	$[d, d, d, a]$	$[a, d, d, d]$
$[a, a, a, a]$	$[a, a, a, a]$	$[a, d, d, d]$

(a) Boundary-driven extreme structure.



Front	Right	Up
$[a, a, a, a]$	$[a, a, a, a]$	$[a, a, a, a]$
$[a, a, a, a]$	$[a, a, a, a]$	$[a, a, a, a]$
$[a, a, a, a]$	$[a, a, a, a]$	$[a, a, a, a]$
$[a, a, a, a]$	$[a, a, a, a]$	$[a, a, a, a]$

(b) Uniform structure.

Figure 4 Conceptual contrast between two $4 \times 4 \times 4$ objects. (a) Depth extremes concentrated along inference boundaries with maximal depth (`d`) dominating interior cells. (b) Uniform shallow configuration with no depth variation. The first configuration provides stronger inferential constraints, while the second admits greater ambiguity in hidden-view reconstruction.

In other words, EQ3D scales deterministically across dimensions, but the *degree of determinacy* depends on how informative the visible structure is. Some objects provide strong symbolic constraints that tightly restrict hidden continuations; others admit multiple logically consistent completions. This distinction between *scalability of reasoning* and *configuration-dependent determinacy* is fundamental to understanding EQ3D's behaviour.

Concavity, Extremal Structure, and Inferential Reach

A key factor affecting coverage is the distribution of qualitative depth values across inference-relevant regions. EQ3D rules propagate information primarily from:

1. **Inference-boundary slices** (e.g., last column of `Right`, top row of `Up`, left column of `Up`, last row of `Right`, last row of `Front`), and
2. **Non-boundary cells** containing holes (`*`) or view-dependent maximum depth symbols.

Visible structure that exhibits strong depth contrasts at these locations tends to produce higher inferential reach. By contrast, configurations dominated by intermediate depth values in non-boundary regions leave more degrees of freedom for hidden views. This structural effect can be illustrated conceptually by contrasting two $4 \times 4 \times 4$ configurations (Figure 4).

In configuration (a), inference-boundary cells contain minimal depth (a), while interior cells contain maximal depth (d). This produces strong extremal cues: propagation across opposite views becomes tightly constrained, and fewer hidden alternatives remain admissible. Table 2 shows substantially higher coverage (66.67%) and a very low interior non-extremal ratio (INER = 0.10), indicating that most interior cells are structurally forced rather than ambiguous.

In configuration (b), uniform depth provides fewer constraints. Although globally consistent, the visible structure does not sufficiently restrict hidden extensions, resulting in lower inferential reach. This is reflected in Table 2 with the reduced coverage (37.50%) and maximal INER (1.00), meaning that all interior cells remain non-extremal and thus weakly informative despite the high boundary-rule share (BSS = 1.00).

Figure 3 shows this behaviour. Object o13 exhibits higher inferential completeness, with broader propagation of depth constraints and lower ambiguity in hidden cells. In contrast, o19 shows reduced coverage and more undetermined entries, reflecting weaker interior constraints and heavier reliance on boundary-driven inference, consistent with the coverage and structural indicators reported in.

Table 2 Coverage and inference indicators for conceptual objects (a) and (b).

Object	Coverage (%)	Collision Rate (CR)	Boundary Share (BSS)	INER
(a) Boundary-driven	66.67	0.3191	0.4255	0.1000
(b) Uniform	37.50	0.1000	1.0000	1.0000

Correlation Patterns and Structural Effects

Correlation analysis between coverage and structural indicators reveals a coherent pattern. The strongest negative correlation with coverage is observed for `V_INTERIOR_OTHER_R` (Spearman $\rho = -0.837$). This metric measures the proportion of visible non-boundary cells whose values are neither holes nor maximum depth. A high value indicates heterogeneous intermediate depth structure. The strong negative correlation confirms that such configurations reduce inferential determinacy. Conversely, coverage correlates positively with indicators of extremal qualitative structure:

- **Maximum-depth ratio (MDR)**: Spearman $\rho = 0.533$,
- **Visible maximum-depth count (V_MAX)**: Spearman $\rho = 0.638$,
- **Combined star/max presence (SAS)**: Spearman $\rho = 0.642$.

These results confirm that holes and maximum-depth symbols act as strong inferential anchors. They reduce the number of admissible hidden completions and increase reconstructability. Boundary Share (BSS) shows a moderate negative correlation with coverage (Spearman $\rho = -0.600$). This suggests that heavy reliance on inference-boundary rules is characteristic of more ambiguous configurations: when interior propagation is weak, the system proportionally depends more on boundary-driven inference. Overall, the correlation profile supports a unified structural interpretation:

Key observation. Coverage increases when visible structure contains strong extremal qualitative cues and decreases when interior regions exhibit intermediate depths that admit multiple consistent completions.

Implications for Qualitative Volumetric Reasoning

These findings highlight a central epistemic property of EQ3D: the framework does not optimise over possible completions, nor does it approximate hidden structure. Instead, it performs principled symbolic propagation grounded in qualitative constraints.

Variation in coverage therefore reflects *intrinsic geometric ambiguity*, not algorithmic deficiency. EQ3D distinguishes between (i) structure that is logically entailed by visible projections, and (ii) structure that remains underdetermined. This distinction is essential in qualitative spatial reasoning, where interpretability and constraint-based inference take precedence over maximal reconstruction.

Limitations. EQ3D operates over discrete qualitative symbols and assumes perfectly aligned orthographic views. It does not currently model uncertainty in the input, such as ambiguous depth assignments or perceptual noise. Depth qualitative values are treated ordinally rather than metrically, and symmetrical configurations may yield multiple qualitatively valid hidden completions.

7 Conclusion and Future Work

This paper presented the EQ3D, a dimension-agnostic extension of the qualitative 3D (Q3D) representation that generalises volumetric reasoning from a fixed $3 \times 3 \times 3$ grid to arbitrary $H \times W \times L$ partitions. EQ3D preserves the symbolic depth vocabulary and the core logical principles of Q3D – depth relativity, volumetric continuity, and view adjacency, while removing dimensional constraints. In doing so, it establishes a uniform reasoning framework capable of handling heterogeneous volumetric resolutions without altering the underlying rule system.

A central contribution of EQ3D lies in the integration of adjacent-view propagation rules with the opposite-view constraints. This combination enables additional deductions from locally shared boundaries and increases inferential reach without introducing procedural complexity. On the other hand, the reasoning mechanism remains fully declarative and size-independent: validation and inference both scale linearly with the number of view cells and hidden targets.

Empirical evaluation across twenty-four structurally diverse objects shows that EQ3D consistently reconstructs hidden views from three orthographic inputs while preserving global cross-view coherence. All inferred cells match the ground truth (100% accuracy), and variation in coverage reflects intrinsic geometric ambiguity rather than algorithmic weakness.

Beyond reconstruction performance, EQ3D provides a principled characterisation of *inferential determinacy*. The EQ3D does not approximate or optimise over candidate completions; instead, it explicitly distinguishes between hidden cells that are logically involved and those that remain underdetermined. This property is a strength of EQ3D as a spatial reasoning model rather than merely a reconstruction technique, since it shows robust explainability.

As future work we intend to explore diverse directions. On one hand, we intend to explore a probabilistic extension of EQ3D, for instance through a ProbLog-based [5] formulation, allowing uncertain or partially reliable input views to be represented with graded confidence. Such an extension would enable reasoning under perceptual noise while preserving symbolic traceability of inference. This would relax the current assumption that the `Front-Right-Up` descriptions are exact inputs. A complementary direction involves developing a neural-symbolic EQ3D approach using a neural front-end which might predict soft `Front-Right-Up` tables from images, point clouds, or depth maps, while EQ3D might enforce structural consistency through logical constraints. This would also provide a possible inference to less regular geometries, since curved, oblique, or non-axis-aligned objects could first be approximated through voxels or part-based abstractions before being translated into qualitative EQ3D descriptors. For that, consistency-aware loss functions might ensure that learned predictions respect adjacency and depth relations, yielding an interpretable hybrid pipeline. On the other hand, recent advances in 3D object detection architectures suggest that the voxel-based and part-aware aggregation mechanisms can effectively capture spatial structure while maintaining computational efficiency [33, 6]. Integrating such representations with symbolic consistency constraints might provide a bridge between geometric learning

and logical reasoning. Moreover, as point-voxel transformers model global and fine geometric structure [22], supporting constraint-aware view prediction, then active-view selection strategies may be explored, where the system might determine which additional orthographic projection would maximally reduce hidden ambiguity.

Taken together, these directions suggest a trajectory in which EQ3D evolves from a purely qualitative inference calculus into a probabilistic, learnable, and perception-aware reasoning module. By maintaining a strong symbolic core while enabling integration with statistical and geometric components, EQ3D offers a promising foundation for interpretable neuro-symbolic reasoning over uncertain inputs and less regular volumetric structures.

References

1 American Dental Association. Dental admission test (dat). https://www.ada.org/en/education/testing/exams/dental-admission-test-dat, 2025. Includes the Perceptual Ability Test (PAT) section.

2 Anthony G. Cohn and Jochen Renz. Chapter 13 qualitative spatial representation and reasoning. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*, pages 551–596. Elsevier, 2008. doi:10.1016/S1574-6526(07)03013-1.

3 DAT Prep. Dat–pat: Perceptual ability test. https://www.dat-prep.com/dat-pat-perceptual-ability, 2025. Accessed: 2026-02-24.

4 Artur d'Avila Garcez, Luis C. Lamb, and Dov M. Gabbay. Neural-symbolic cognitive reasoning. *Springer Handbook of Computational Intelligence*, 2015.

5 Luc De Raedt, Angelika Kimmig, and Hannu Toivonen. Problog: a probabilistic prolog and its application in link discovery. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, IJCAI'07, pages 2468–2473, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc. URL: https://dl.acm.org/doi/10.5555/1625275.1625673.

6 Jiajun Deng, Shaoshuai Shi, Peiwei Li, Wengang Zhou, Yanyong Zhang, and Houqiang Li. Voxel r-cnn: Towards high performance voxel-based 3d object detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(2):1201–1209, May 2021. doi:10.1609/aaai.v35i2.16207.

7 Frank Dylla, Jae Hee Lee, Till Mossakowski, Thomas Schneider, André van Delden, Jasper van de Ven, and Diedrich Wolter. A survey of qualitative spatial and temporal calculi: Algebraic and computational properties. *ACM Comput. Surv.*, 50(1):7:1–7:39, 2017. doi:10.1145/3038927.

8 Martí Ejarque and Zoe Falomir. Ejarque24/Extended-Q3D. Software (visited on 2026-07-20). URL: https://github.com/Ejarque24/Extended-Q3D.

9 Z. Falomir. Towards a qualitative descriptor for paper folding reasoning. In *Proc. of the 29th International Workshop on Qualitative Reasoning*, 2016. Co-located with IJCAI'2016 in New York, USA.

10 Zoe FAlomir. A Qualitative Model for Reasoning about 3D Objects using Depth and Different Perspectives. In Tomasz Lechowski, Przemyslaw Andrzej Walega, and Michal Zawidzki, editors, *Proceedings of the 2015 Workshop on Logics for Qualitative Modelling and Reasoning (LQMR@FedCSIS 2015), Lódz, Poland, September 13-16, 2015*, volume 7 of *Annals of Computer Science and Information Systems*, pages 3–11, 2015. URL: https://annals-csis.org/Volume_7/drp/370.html.

11 Zoe Falomir. A qualitative model to reason about object rotations (QOR) applied to solve the cube comparison test (CCT). *ArXiv preprint arXiv.2601.08382*, 2025. doi:10.48550/arXiv.2601.08382.

12 Zoe Falomir, Albert Pich, and Vicent Costa. Spatial reasoning about qualitative shape compositions. *Ann. Math. Artif. Intell.*, 88(5-6):589–621, 2020. doi:10.1007/S10472-019-09637-7.

13 Zoe Falomir, Ruben Tarin, Aurelio Puerta, and Pablo Garcia-Segarra. An interactive game for training reasoning about paper folding. *Multim. Tools Appl.*, 80(5):6535–6566, 2021. doi:10.1007/s11042-020-09830-5.

14 Kenneth D. Forbus. Qualitative modeling. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(4):374–391, 2011. doi:10.1002/wcs.115.

15 Christian Freksa. *Qualitative Spatial Reasoning*, pages 361–372. Springer Netherlands, Dordrecht, 1991. doi:10.1007/978-94-011-2606-9_20.

16 Jinn-Kwei Guo, Chin-Hsing Chen, and Jiann-Der Lee. Multi-polyhedron reconstruction in a three-view system using relaxation. *Signal Processing*, 77(2):171–193, 1999. doi:10.1016/S0165-1684(99)00031-6.

17 Hamid Haghshenas Gorgani, Alireza Jahantigh, and Sadegh Sadeghi. 3d model reconstruction from two orthographic views using fuzzy surface analysis. *European Journal of Sustainable Development Research*, 3, March 2019. doi:10.29333/ejosdr/5726.

18 Wenyu Han, Siyuan Xiang, Chenhui Liu, Ruoyu Wang, and Chen Feng. Spare3d: A dataset for spatial reasoning on three-view line drawings. In *CVPR*, pages 14678–14687, June 2020. doi:10.1109/CVPR42600.2020.01470.

19 Long Hoang. 3d solid reconstruction from 2d orthographic views. In Branislav Sobota and Dragan Cvetković, editors, *Mixed Reality and Three-Dimensional Computer Graphics*, chapter 6. IntechOpen, London, 2020. doi:10.5772/intechopen.91977.

20 Wentao Hu, Jia Zheng, Zixin Zhang, Xiaojun Yuan, Jian Yin, and Zihan Zhou. Plankassembly: Robust 3d reconstruction from three orthographic views with learnt shape programs. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 18449–18459, 2023. doi:10.1109/ICCV51070.2023.01695.

21 David A. Huffman. Impossible objects as nonsense sentences. In *Machine Intelligence 6*, 2012.

22 Guangyu Ji, Jun Lu, Chengtao Cai, and Kaibin Qin. A point-voxel transformer for point cloud object detection with spatial and channel attention. *Engineering Applications of Artificial Intelligence*, 163:113141, 2026. doi:10.1016/j.engappai.2025.113141.

23 Lefteris M. Kirousis and Christos H. Papadimitriou. The complexity of recognizing polyhedral scenes. *Journal of Computer and System Sciences*, 37(1):14–38, 1988. doi:10.1016/0022-0000(88)90043-8.

24 Nikhil Krishnaswamy, Scott Friedman, and James Pustejovsky. Combining deep learning and qualitative spatial reasoning to learn complex structures from sparse examples with noise. In *Proc. of 33rd AAAI Conf. on Artificial Intelligence and 31st Innovative Applications of Artificial Intelligence Conference and 9th AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI'19/IAAI'19/EAAI'19. AAAI Press, 2019. doi:10.1609/aaai.v33i01.33012911.

25 G. Ligozat. *Qualitative Spatial and Temporal Reasoning*. MIT Press, Wiley-ISTE, London, 2011.

26 Jianzhuang Liu, Liangliang Cao, Zhenguo Li, and Xiaoou Tang. Plane-based optimization for 3d object reconstruction from single line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):315–327, 2008. doi:10.1109/TPAMI.2007.1172.

27 A. Lovett, M. Dehghani, and K. Forbus. Learning of qualitative descriptions for sketch recognition. In *Proc. 20th Int. Workshop on Qualitative Reasoning (QR), Hanover, USA, July*, 2006.

28 A. Lovett and K. Forbus. Cultural commonalities and differences in spatial problem-solving: A computational analysis. *Cognition*, 121(2):281–287, 2011. doi:10.1016/j.cognition.2011.06.012.

29 A. Lovett and K. Forbus. Modeling visual problem solving as analogical reasoning. *Psychological Review*, 124(1):60–90, 2017.

30 David Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. The MIT Press, July 2010. doi:10.7551/mitpress/9780262514620.001.0001.

31 Albert Pich and Zoe Falomir. Logical composition of qualitative shapes applied to solve spatial reasoning tests. *Cogn. Syst. Res.*, 52:82–102, 2018. doi:10.1016/J.COGSYS.2018.06.002.

32 Jochen Renz and Bernhard Nebel. On the complexity of qualitative spatial reasoning: a maximal tractable fragment of the region connection calculus. *Artif. Intell.*, 108(1–2):69–123, March 1999. doi:10.1016/S0004-3702(99)00002-8.

33 Shaoshuai Shi, Zhe Wang, Jianping Shi, Xiaogang Wang, and Hongsheng Li. From points to parts: 3d object detection from point cloud with part-aware and part-aggregation network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(8):2647–2664, 2021. doi:10.1109/TPAMI.2020.2977026.

34 Sheryl Sorby. *Developing Spatial Thinking*. Higher Education Services, 2016.

35 Ken Tomiyama and Ken'ichi Nakaniwa. Reconstruction of 3D solid model from three orthographic views — Top-down approach. In Rangachar Kasturi and Karl Tombre, editors, *Graphics Recognition Methods and Applications*, pages 260–269, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.

36 Jan Wielemaker, Tom Schrijvers, Markus Triska, and Torbjörn Lager. SWI-Prolog. *Theory and Practice of Logic Programming (TPLP)*, 12(1-2):67–96, 2012. doi:10.1017/S1471068411000494.

37 Falomir Zoe and Oliver Eric. Q3D-Game: A Tool for Training Users' 3D Spatial Skills. In Paulo Novais and Shin'ichi Konomi, editors, *Intelligent Environments 2016 - Workshop Proc. of 12th Int. Conference on Intelligent Environments, IE 2016, London, UK, September 14-16, 2016*, volume 21 of *Ambient Intelligence and Smart Environments*, pages 187–196. IOS Press, 2016. doi:10.3233/978-1-61499-690-3-187.

38 Changqing Zou, Shifeng Chen, Hongbo Fu, and Jianzhuang Liu. Progressive 3d reconstruction of planar-faced manifold objects with drf-based line drawing decomposition. *IEEE Transactions on Visualization and Computer Graphics*, 21(2):252–263, 2015. doi:10.1109/TVCG.2014.2354039.