

# The Space of Possible Questions A Typology of Interrogatives and a Totality Theorem for Question-Handling Systems

J. Konstapel — Leiden, The Netherlands. June 11 2026.

## Abstract

Automated question-answering systems routinely fail not because they retrieve poor answers, but because they recognize only one kind of question: the kind their sources can answer. This paper assembles the formal, linguistic, and cognitive literature on questions into a single typology and proves two complementary results. First, an *impossibility result*: no system can be complete over the **content** of the question space, since the space of askable questions is not recursively closeable — over any ordering of questions, a further well-formed question about that ordering can be posed. Second, a *possibility result*: a system can be complete over the **treatment** of questions. We show that four binary tests, each total over well-formed interrogatives, induce a partition of the question space into eight mutually exclusive, jointly exhaustive disposal classes, four of which are principled non-answers (referral, presupposition correction, deferred monitoring, and perspectival display). A question-handling system implementing all eight classes is a total function over the question space; a system implementing fewer must guess outside its implemented classes, with empirically observable failure rates. The result reframes the design target of civic and institutional question systems: not maximal answering, but total honest disposal.

## 1. Introduction

Recent evaluations of municipal chatbots in the Netherlands found correct responses to roughly one in ten citizen questions, with the majority of failures taking the form of confident but wrong answers triggered by keyword matching. We argue that this failure mode is structural rather than incidental. Such systems implement a single disposal — retrieve-and-answer — and therefore have no licit response available for the large classes of questions that cannot, in principle or in practice, be answered from their sources: future contingents, value questions, questions resting on false presuppositions, questions belonging to another institution's mandate, and utterances that are not information requests at all. Lacking a defined treatment for these classes, the system guesses.

The remedy requires an answer to a prior theoretical question: *what kinds of questions are possible, and can the kinds be enumerated?* Section 2 reviews four formal traditions that converge on what a question is. Section 3 organizes the established typological dimensions. Section 4 states and proves the impossibility of content-completeness. Section 5 states and proves the possibility of treatment-completeness — the central result. Section 6 applies the framework to its own founding question. Section 7 draws design implications. An annotated bibliography follows.

## 2. What a Question Is: Four Convergent Formalizations

Four research traditions, largely independent in origin, characterize the question in compatible ways.

**(i) Answer-set semantics.** Hamblin (1958) proposed that a question denotes the set of its possible answers; to understand a question is to know what would count as answering it. Belnap and Steel (1976) developed this into the first systematic erotetic logic, with a formal apparatus for direct answers, presuppositions, and question equivalence.

**(ii) Partition semantics.** Groenendijk and Stokhof (1984) analyze a question as a partition of logical space: the question divides the set of possible worlds into cells, and a complete answer locates the actual world in exactly one cell. Inquisitive semantics (Ciardelli, Groenendijk and Roelofsen, 2018) generalizes the account, treating informative and inquisitive content as two components of a single semantic object.

**(iii) Interrogative epistemology.** Hintikka's interrogative model of inquiry treats questions as moves in a knowledge-seeking game against a source; questioning and investigating are formally the same activity. This recovers, in modern dress, the etymological core of the Germanic *fr̥gōnq* and the Latin *quaestio*, both of which denoted investigation before they denoted speech.

**(iv) Cognitive theory.** In Schank's account of dynamic memory, a question is an explicitly reported expectation failure: the questioner announces where a knowledge structure broke. Bromberger (1992) similarly characterizes the questioner as occupying a "p-predicament" — knowing enough to pose the question while lacking what would resolve it.

**Synthesis.** A question is a *self-reported partition uncertainty*: a declaration by an agent of which cells of logical space it cannot currently distinguish. Answering is refining the partition to a single cell; the literature's apparent diversity concerns the formal vehicle, not the underlying object. This synthesis grounds the typology below: questions can differ in (a) the linguistic form of the declaration, (b) the kind of cell-structure declared, (c) the epistemic state and purpose of the declarer, and (d) what can lawfully be done about the declared uncertainty. These are the four axes of Section 3 and the basis of the theorem in Section 5.

## 3. The Dimensions of Question Space

### 3.1 Form (closed inventory)

The syntactic inventory of interrogatives is finite and stable across the well-studied languages: polar (yes/no), alternative, constituent (*wh-*) questions, quantity questions, embedded interrogatives, tag and echo questions, rhetorical questions, exam questions (where the asker knows the answer and tests the answerer), and loaded questions (carrying a contestable presupposition; Belnap and Steel 1976, §2). The *wh*-paradigm is itself a closed class whose members each constrain the semantic type of admissible answers (person/agent, entity, place, time, cause or justification, manner or procedure, selection, quantity). No new interrogative forms are being added to natural languages; the form axis is therefore exhaustively enumerable.

### 3.2 Epistemic type of the sought knowledge

Synthesizing the taxonomies of Graesser and Person (1994), Li and Roth (2002), Dillon (1984), and the classical four-causes analysis, the knowledge sought by a question falls into thirteen types, grouped by decidability:

- *Decidable in principle (past/present)*: fact or state; event; definition or identity; quantity; location; mechanistic cause; agent intention (decidable only via testimony); procedure.
- *Undecidable now (future-directed)*: prediction; counterfactual or hypothetical.

- *Outside the fact-space*: evaluation; norm or prescription.
- *Reflexive*: meta-questions, including questions about the typology itself.

The decisive property of this axis is that **epistemic type determines what an answer can be**. Fact-type questions close with a sourced statement; future contingents cannot close now but can be monitored until the world answers; evaluative and normative questions are not closed by facts at all; meta-questions close against the system's own specification.

### 3.3 The asker's state

Taylor (1968) showed that an expressed question is the fourth and most degraded form of an information need (visceral → conscious → formalized → compromised); the question as received has already been adapted to what the asker believes the system can handle. Belkin (1980) formalized the consequence as the Anomalous State of Knowledge: the asker cannot, by definition, fully specify what is missing. Two design corollaries follow: a system must sometimes negotiate one level upward (clarification), and answers should be permitted to exceed the literal question while marking the excess. A further, empirically documented property of this axis is *non-asking*: the propensity to question is unevenly distributed across language proficiency, institutional trust, and self-efficacy, so any question stream is a selective sample of the underlying uncertainty distribution.

### 3.4 Pragmatic intent

Following speech-act theory (Searle 1969), an interrogative utterance may serve distinct illocutionary purposes: to know, to decide, to act, to verify, to influence, to affiliate, or to vent. Intent does not change the semantics of the question but changes what counts as a satisfactory disposal of it — a point developed in Section 5.

## 4. The Impossibility Result: No Content-Completeness

**Proposition 1.** *There is no finite or recursively enumerable catalogue of questions of which every possible question is a member.*

*Argument.* (a) The constituent-question schema is productive: for every nameable entity, property, or proposition, a well-formed question can be generated; the supply of nameables is unbounded. (b) Erotetic inference (Wiśniewski 1995) generates new questions from existing questions and declaratives, without fixpoint. (c) Diagonally: let  $\mathbf{O}$  be any proposed complete ordering of questions; the question "does  $\mathbf{O}$  contain every question, including this one?" is well-formed, is *about*  $\mathbf{O}$ , and its admission generates a successor question about the amended ordering, ad infinitum. Hence the content of question space is not closeable. ■

Proposition 1 explains why "a system that can answer every question" is, read naively, a contradiction: the demand quantifies over an unclosed domain. Any tenable completeness claim must therefore quantify over something else.

## 5. The Possibility Result: Treatment-Completeness

### 5.1 Four total binary tests

Define, for any well-formed interrogative utterance  $Q$  received by a system  $S$  with source set  $\Sigma$ :

- **T1 (Request test):** is  $Q$  an information request? (Rhetorical, venting, abusive, and nonsensical utterances fail T1. The test is total: every utterance either does or does not solicit information, with classification — not class membership — being the only locus of error.)
- **T2 (Presupposition test):** do the presuppositions of  $Q$  hold, so far as  $\Sigma$  can determine? (Belnap and Steel's apparatus makes presupposition a formal property of the question, not of the asker.)
- **T3 (Decidability test):** is the epistemic type of  $Q$  decidable in principle from facts about the past or present? (Future contingents and evaluative/normative questions fail T3 — for different reasons, distinguished below.)
- **T4 (Source test):** if decidable in principle, is  $Q$  decidable now from  $\Sigma$ , from reliable general knowledge, from another institution's holdings, or from none of these without clarification?

Each test is total over well-formed interrogatives: it assigns every question to exactly one branch.

## 5.2 The eight disposal classes

The tests induce the following partition:

Class	Path	Disposal
<b>H1</b> Locally	T1✓ T2✓ T3✓ T4:	sourced answer; where sources diverge from official
<b>H2</b> Generally answerable	T1✓ T2✓ T3✓ T4: general	answer, explicitly marked as non-local
<b>H3</b> Held	T1✓ T2✓ T3✓ T4:	directed referral
<b>H4</b> Underdetermin	T1✓ T2✓ T3✓ T4: clarification needed	exactly one negotiation move (Taylor), then re-entry
<b>H5</b> Presupposition	T1✓ T2✗	the presupposition is named and tested before any answer
<b>H6</b> Future-	T1✓ T2✓ T3✗	report what is currently fixed; offer monitoring until the
<b>H7</b> Value question	T1✓ T2✓ T3✗ (normative)	display the existing positions evenhandedly; offer connection among co-holders of the question
<b>H8</b> Non-	T1✗	acknowledge, reformulate, or decline, as appropriate;

**Theorem (Totality).** *A system implementing all eight disposal classes is a total function over the space of well-formed interrogative utterances: every question receives exactly one defined disposal.*

*Proof.* Exhaustiveness: T1–T4 are total and the eight classes are the leaves of their decision tree; every question traverses the tree to exactly one leaf. Exclusivity: the leaves are pairwise disjoint by construction (each pair differs in at least one test outcome). Hence the disposal assignment is a total function. ■

**Corollary.** *A system implementing  $k < 8$  classes possesses no licit response for questions falling in the  $8 - k$  unimplemented classes and must therefore either refuse indiscriminately or guess. The empirical failure profile of single-class (retrieve-and-answer) systems — confident error on future contingents, loaded questions, and out-of-mandate questions — is the corollary made visible.*

## 5.3 Two clarifications of the claim's strength

First, the theorem concerns *class existence*, not *classification accuracy*: a deployed system may misassign a question to the wrong class, but no question lacks a class. Misclassification is an

engineering error rate to be measured against an annotated calibration set; classlessness, the structural defect of current systems, is eliminated. Second, four of the eight disposals are principled non-answers. The theorem's force lies precisely here: referral (H3), presupposition correction (H5), deferred monitoring (H6), and evenhanded perspectival display with affiliation (H7) are *truthful treatments*, and a system honest in these classes dominates, in expected veracity, any system that converts them into guessed answers.

## 6. Self-Application

The framework must dispose of the question that motivated it: "*Do we now know everything there is to know about asking questions?*" Classification: polar form; meta epistemic type; verificative intent. The question carries the presupposition that "everything about questions" denotes a closeable content set; by Proposition 1 the presupposition fails, routing the question through H5 to H1. The corrected, sourced answer: *no* with respect to content, which is provably unclosed; *yes* with respect to treatment, which the Totality Theorem closes. The framework thereby answers the question of its own completeness without stepping outside itself — the strongest form of completeness that Proposition 1 permits.

## 7. Implications for Question-Handling Systems

(1) The design target of civic and institutional question systems should be total honest disposal, not maximal answering; the eight classes constitute the specification. (2) Classes H6 and H7 define product capabilities absent from current systems: persistent monitoring of future contingents ("the question keeps watch") and affiliation among co-holders of value questions. (3) Because classification is the single point of engineering error, the classifier must be treated as a measurement instrument: versioned prompts and models, an annotated calibration set, and regression gates on every change. (4) The question stream is a selective sample (§3.3) and is more revealing than the answer stream; aggregates may circulate, individual questions must not.

## 8. Annotated References

**Hamblin, C. L. (1958).** "Questions." *Australasian Journal of Philosophy* 36(3), 159–168. The founding postulates of modern erotetics: a question is known through its answer set; an answer is a statement that resolves it. The origin of answer-set semantics, presupposed by all later formal work.

**Belnap, N. D., & Steel, T. B. (1976).** *The Logic of Questions and Answers*. Yale University Press. The first systematic erotetic logic. Supplies the formal machinery used here for direct answers, question presuppositions (the basis of test T2 and class H5), and question equivalence.

**Åqvist, L. (1965).** *A New Approach to the Logical Theory of Interrogatives*. Uppsala. Reduces questions to epistemic imperatives ("bring it about that I know..."), anticipating the treatment of questions as requests for epistemic state change — the reading under which test T1 is well-defined.

**Hintikka, J. (1999).** *Inquiry as Inquiry: A Logic of Scientific Discovery*. Kluwer. The mature statement of the interrogative model: inquiry as a questioning game against sources. Grounds the identification of questioning with investigation and hence the legitimacy of disposal classes that investigate rather than merely retrieve.

**Groenendijk, J., & Stokhof, M. (1984).** *Studies on the Semantics of Questions and the Pragmatics of Answers*. Dissertation, University of Amsterdam. Partition semantics: a question

partitions logical space; complete answers locate the world in one cell. Supplies the paper's core semantic object (the partition) and the criterion of closure used in §7.

**Ciardelli, I., Groenendijk, J., & Roelofsen, F. (2018).** *Inquisitive Semantics*. Oxford University Press. The current standard unification of informative and inquisitive content. Demonstrates that the question/statement boundary is itself semantic, supporting the totality of test T1.

**Wiśniewski, A. (1995).** *The Posing of Questions: Logical Foundations of Erotetic Inferences*. Kluwer. Formalizes how questions arise from questions and declaratives (erotetic inference). Provides limb (b) of Proposition 1: question space is generatively open.

**Bromberger, S. (1992).** *On What We Know We Don't Know*. University of Chicago Press. Analyzes why-questions and the questioner's "p-predicament." Clarifies the asker-side condition under which a question is sincere — knowing enough to ask, lacking what resolves — used in §2(iv).

**Searle, J. R. (1969).** *Speech Acts*. Cambridge University Press. Classifies questions among directives and analyzes indirect speech acts. Grounds the pragmatic-intent axis (§3.4) and the recognition that not every interrogative utterance is an information request (test T1).

**Taylor, R. S. (1968).** "Question-Negotiation and Information Seeking in Libraries." *College & Research Libraries* 29(3), 178–194. The four levels of information need and the negotiation of the "compromised" question. The empirical basis of class H4 and of the one-clarification rule.

**Belkin, N. J. (1980).** "Anomalous States of Knowledge as a Basis for Information Retrieval." *Canadian Journal of Information Science* 5, 133–143. The asker cannot fully specify the gap being asked about. Justifies marked answer-broadening beyond the literal question.

**Dillon, J. T. (1984).** "The Classification of Research Questions." *Review of Educational Research* 54(3), 327–361. A disciplined survey of question classifications across fields; the methodological caution against ad hoc taxonomies that the present four-test derivation is designed to meet.

**Graesser, A. C., & Person, N. K. (1994).** "Question Asking During Tutoring." *American Educational Research Journal* 31(1), 104–137. The eighteen-category question taxonomy standard in education and cognitive science; the empirical backbone of the thirteen epistemic types in §3.2.

**Li, X., & Roth, D. (2002).** "Learning Question Classifiers." *Proceedings of COLING 2002*. The TREC question-classification taxonomy (6 coarse, 50 fine classes) on which automated QA classification rests; demonstrates that epistemic-type classification is machine-learnable — the feasibility premise of §7(3).

**Schank, R. C. (1982).** *Dynamic Memory*. Cambridge University Press; and (1986). *Explanation Patterns*. Erlbaum. Understanding as question-driven; the question as reported expectation failure. Supplies the cognitive reading in §2(iv) and the rationale for treating received questions as the most informative events a learning system can ingest.

**Collingwood, R. G. (1940).** *An Essay on Metaphysics*. Oxford University Press. The logic of question and answer: every statement is an answer to a question, and meaning is relative to the question answered. The philosophical warrant for refusing to evaluate answers apart from the questions that elicited them.

**Bloom, B. S. (Ed.) (1956). *Taxonomy of Educational Objectives*. Longmans.** The canonical hierarchy of cognitive demand behind questions; included as the historical template — and cautionary example — of a content-based question hierarchy, superseded here by treatment-based classification.

*Correspondence: constable.blog. The typology carries its own falsification criterion: any well-formed interrogative utterance demonstrably assignable to none of the eight disposal classes refutes the Totality Theorem and forces a ninth class.*