International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Knowledge-Grounded Multi-Agent Architecture for Vedic Astrological Question Answering: Operationalizing Bṛhat Parāśara Horā Śāstra across 1,000 Specialized Agents

Pranav Anand

Jothishi Online and Advertising Services Private Limited (RetroGrade), Bengaluru, Karnataka, India

Abstract

Vedic astrology (Jyotiṣa) encodes one of the largest continuously practiced interpretive knowledge systems in human history, yet its analytical procedures remain largely undigitized in a form suitable for open-ended natural-language interaction. We present the knowledge architecture underlying RetroGrade, a conversational platform that maps free-text life questions to structured Parāśarī analysis. The system decomposes the space of user intents into a taxonomy of approximately 14,000 probable questions spanning the principal domains of human concern — career, marriage and relationships, wealth, health, education, progeny, property, litigation, travel, spirituality, personality and event timing — and services them through a federation of 1,000 specialized agents. Each agent is grounded in the rule corpus of Bṛhat Parāśara Horā Śāstra (BPHS) and allied classical texts, and is scoped to a coherent analytical procedure that invokes the appropriate computational primitives: the sixteen divisional charts (ṣoḍaśavarga), the twenty-seven nakṣatras, planetary states (avasthā), multiple daśā systems, and the six-fold strength model (ṣaḍbala), including the motional strength (ceṣṭā-bala) from which retrograde planets derive their prominence and from which the platform takes its name. A two-stage semantic router maps each question to its agent(s) through dense-embedding retrieval over the taxonomy, and a grounded generation layer composes answers strictly from computed chart facts and retrieved rules under explicit guardrails. We describe the taxonomy, the agent-specification schema, the routing pipeline and the computational engine; we propose an evaluation protocol for routing accuracy and rule-faithfulness; and we position the contribution explicitly as an exercise in knowledge engineering and computational modeling of a classical tradition, rather than as a claim about the empirical predictive validity of astrology.

Keywords: Vedic astrology; Jyotiṣa; knowledge representation; multi-agent systems; semantic routing; retrieval-augmented generation; conversational AI; Bṛhat Parāśara Horā Śāstra; divisional charts; computational humanities.

1. Introduction

Jyotiṣa, the classical astral science of the Indian subcontinent, is among the most elaborate rule-based interpretive systems ever formalized. Its canonical treatise, *Bṛhat Parāśara Horā Śāstra* (BPHS), together with allied works such as *Sārāvalī*, *Phaladeepikā*, *Bṛhat Jātaka* and the *Jaimini Sūtras*, specifies thousands of conditional rules that map the geometric configuration of the sky at a moment of birth onto statements about a person's life. The system is procedural in a way that is unusual among divinatory traditions: it prescribes explicit calculations (planetary longitudes, divisional subdivisions, strength scores, time-period sequences) and explicit combinatorial rules over their outputs. This procedural character is precisely what makes it amenable to computational modeling.

Despite this, contemporary software for Jyotiṣa has remained overwhelmingly *calculational* rather than *interpretive*. Standard packages compute charts, daśā timelines and strength tables with high numerical fidelity, but they leave the interpretive step — the mapping from a specific human question to the relevant configuration and its classical reading — to a human astrologer. The consumer who types "Will I get a government job, and when?" or "Why do my relationships keep ending the same way?" cannot be served by a chart image alone. Bridging that gap at scale, across the full breadth of questions people actually ask, is an open problem in applied knowledge engineering.

This paper describes how the RetroGrade platform addresses that problem. Rather than training a single monolithic model to "do astrology," we treat the domain as a large but enumerable space of analytical procedures and decompose it explicitly. Our contributions are:

1. **A life-question taxonomy** of roughly 14,000 probable user questions, organized into thirteen top-level life domains and a hierarchy of sub-intents, constructed to give near-exhaustive coverage of the questions a person is likely to bring to a consultation.
2. **A federation of 1,000 knowledge-grounded agents**, each scoped to a coherent BPHS-derived analytical procedure and each declaring the computational primitives (vargas, nakṣatras, avasthās, daśās, ṣaḍbala components) it consults.
3. **A two-stage semantic router** that maps an arbitrary free-text question to the correct agent or small set of agents via dense-embedding retrieval, avoiding a brittle keyword taxonomy and an expensive per-query classifier LLM call.
4. **A computational Jyotiṣa engine** that supplies each agent with the exact chart facts its rules require, and **a grounded synthesis layer** that composes the final answer only from computed facts and retrieved rules, under guardrails appropriate to a sensitive advisory domain.

We also state plainly what this paper is *not*. It is not an empirical test of whether astrological predictions are accurate. Section 13 makes our epistemological position explicit: the scientific object here is the faithful digitization and operationalization of a classical interpretive system and the natural-language-processing architecture that makes it usable, not the truth-value of the tradition's claims about the world.

2. Background: The Computational Primitives of Parāśarī Jyotiṣa

This section summarizes the classical constructs that our agents operate over. Each is a well-defined, computable object; together they form the "feature space" of the system. Readers familiar with Jyotiṣa may skip to Section 3.

2.1 The birth chart and the sixteen divisional charts (Ṣoḍaśavarga)

The base object is the *rāśi* chart (D-1): the twelve zodiacal signs, the twelve houses (*bhāva*) reckoned from the ascendant (*lagna*), and the placements of the nine grahas — the Sun, Moon, Mars, Mercury, Jupiter, Venus and Saturn, plus the two lunar nodes Rāhu and Ketu. Beyond the D-1, BPHS defines a set of *divisional* or *varga* charts, each produced by subdividing every sign into a fixed number of parts and re-mapping those parts onto signs. Each varga is a magnifying lens on a particular department of life. The sixteen-fold scheme (ṣoḍaśavarga) is summarized in Table 1.

Chart	Division	Primary domain of signification
D-1 Rāśi	1	The whole life; body, environment, overall template
D-2 Horā	2	Wealth, resources, sustenance
D-3 Drekkāṇa	3	Siblings, courage, initiative
D-4 Chaturthāṁśa	4	Property, fixed assets, inner happiness
D-7 Saptāṁśa	7	Children, progeny, creative lineage
D-9 Navāṁśa	9	Spouse, dharma, and the strength of every planet
D-10 Daśāṁśa	10	Career, profession, public standing
D-12 Dvādaśāṁśa	12	Parents and ancestry
D-16 Ṣoḍaśāṁśa	16	Vehicles, luxuries, material comforts
D-20 Viṁśāṁśa	20	Spiritual practice and progress
D-24 Chaturviṁśāṁśa	24	Education, learning, scholarship
D-27 Bhāṁśa	27	Innate strengths and weaknesses
D-30 Triṁśāṁśa	30	Misfortunes, adversity, hidden risks
D-40 Khavedāṁśa	40	Maternal legacy; auspicious/inauspicious effects
D-45 Akṣavedāṁśa	45	Paternal legacy; conduct and character
D-60 Ṣaṣṭyāṁśa	60	Fine-grained karmic residue; final arbiter

Table 1. The sixteen divisional charts (ṣoḍaśavarga) and their principal significations. The Navāṁśa (D-9) additionally functions as a universal strength check applied across all other analyses.

The interpretive principle that makes vargas powerful is *cross-confirmation*: a promise seen in the D-1 is considered reliable only if it is corroborated in the relevant divisional chart. A career claim, for example, is weighed in the D-1 and the D-10 jointly; a marriage claim in the D-1 and the D-9. Our agents encode exactly which vargas are authoritative for their question and require corroboration before a strong statement is generated.

2.2 Nakṣatras and the lunar framework

Superimposed on the twelve signs are the twenty-seven *nakṣatras* (lunar mansions), each spanning 13°20′ and each subdivided into four *pādas* (quarters), giving 108 pādas that map onto the Navāṁśa. Every nakṣatra has a ruling planet, a presiding deity, a *gaṇa* (temperament class) and a symbolic image; together these encode a finer layer of personality, motivation and compatibility than the signs alone. Critically, the nakṣatra of the natal Moon is the seed of the most widely used time-period system (Section 2.4). Our personality, compatibility and timing agents read the chart at nakṣatra and pāda resolution rather than sign resolution.

2.3 Planetary states (Avasthā)

BPHS describes several schemes of *avasthā* — the "state" or "condition" a planet is in — which modulate how fully and how benevolently it can deliver its results. These states are computable from position and association and are treated by our engine as modifiers on any rule that involves the planet. The principal schemes are summarized in Table 2.

Avasthā scheme	States	Determined by	Effect on interpretation
Bālādi (5)	infant, adolescent, adult, old, dead	degree within the sign	how much of the planet's result manifests
Jāgradādi (3)	awake, dreaming, sleeping	dignity (exalted / own / debilitated)	alertness and effectiveness of the result
Dīptādi (9)	e.g. bright, content, delighted, agitated	dignity, combustion, war, aspects	quality and comfort of the result
Lajjitādi (6)	ashamed, proud, hungry, thirsty, delighted, agitated	houses and associations (nodes, malefics, benefics)	emotional colouring of the result

Table 2. Principal avasthā (planetary-state) schemes used as interpretive modifiers. A planet's raw significations are scaled and coloured by its avasthā before any rule fires.

2.4 Time-period systems (Daśā)

A birth chart states *what* may happen; the *daśā* systems state *when*. A daśā is a sequence of planetary periods and sub-periods that partition the life, during which the ruling planet's promises tend to fructify. Multiple daśā systems coexist and are selected by conditional applicability rules; the platform computes several and reconciles them for timing questions (Table 3).

Daśā system	Cycle	Basis	Typical use
Viṁśottarī	120 years	natal Moon's nakṣatra	primary, general-purpose timing
Aṣṭottarī	108 years	conditional (nakṣatra/lagna)	corroborating timing under conditions
Yoginī	36 years	natal Moon's nakṣatra	quick timing and event checks
Kālachakra	variable	nakṣatra pāda → sign sequence	spiritual and major-event timing
Chara (Jaimini)	sign-based	movable-sign counting	event timing in the Jaimini system

Table 3. Daśā systems computed by the engine. Timing agents run Viṁśottarī by default and reconcile it with Yoginī and, where applicable, Chara daśā, reporting agreement across systems as a confidence signal.

Each daśā period is nested — mahādaśā (major), antardaśā (sub), pratyantardaśā (sub-sub) and finer — so that timing can be resolved from decades down to weeks. Our timing agents traverse this nesting and cross-check the daśā lord's condition (its dignity, avasthā and ṣaḍbala) before asserting that a period is favourable.

2.5 The six-fold strength model (Ṣaḍbala)

Whether a planet can deliver on its promise depends on its *strength*. BPHS quantifies this through *ṣaḍbala*, the six-fold strength, in which six independent components are computed in a common unit (rūpa / ṣaṣṭyāṁśa) and summed. This turns “is this planet strong?” into an arithmetic question, and it is central to how our agents move from possibility to likelihood. The six components are listed in Table 4.

Component	What it measures
Sthāna-bala (positional)	strength from sign dignity, divisional placement and odd/even placement
Dig-bala (directional)	strength from the planet occupying its preferred cardinal direction
Kāla-bala (temporal)	strength from day/night, lunar phase, year/month/day/hour lordship and planetary war
Ceṣṭā-bala (motional)	strength from the planet's apparent motion — maximal near retrogression
Naisargika-bala (natural)	the fixed, intrinsic brightness ranking of the grahas
Dṛk-bala (aspectual)	net strength contributed by benefic and malefic aspects onto the planet

Table 4. The six components of ṣaḍbala. Each is computed in a common unit and summed; agents compare the total against classical minima and against the other planets in the same chart.

2.6 Retrograde motion and Ceṣṭā-bala — the system's namesake

Of the six strength components, *ceṣṭā-bala* — motional strength — is the one most directly tied to the phenomenon the platform is named for. A planet is *vakrī* (retrograde) when, viewed from the geocentric frame of the birth chart, it appears to move backwards against the fixed stars. No planet truly reverses; the appearance arises when the Earth and the planet, orbiting the Sun at different speeds, overtake one another. At the moment of stationing into retrogression a planet makes its closest approach to Earth and shines at its brightest, and BPHS accordingly awards it near-maximal ceṣṭā-bala. Retrogression is therefore not a mark of weakness but a computable surge of strength, turned inward.

This has concrete interpretive consequences that several classical authorities discuss and that our agents encode. *Sārāvalī* holds that retrograde benefics become notably powerful while retrograde malefics grow restless; *Phaladeepikā* states that a retrograde planet can behave as though exalted even in an inimical sign; and the widely used *nīchābhilāṣī / uchchābhilāṣī* principle holds that a debilitated planet in retrogression can act as if exalted, and an exalted planet in retrogression can lose part of its ease. Because the classical authorities are not unanimous, our agents treat retrogression as an *intensifier whose sign the surrounding chart determines*, rather than as a fixed positive or negative flag — and they surface the disagreement transparently rather than adjudicating it silently. The engine computes vakrī status and ceṣṭā-bala for all five true planets (the nodes are always retrograde by nature and are read differently), making retrograde analysis a first-class, always-available signal across the whole agent federation.

3. Related Work

Our work sits at the intersection of three lines of research. First, **computational treatments of astrology and other rule-based traditional systems** have historically taken the form of calculational software and small rule-based expert systems; these achieve high numerical accuracy on charts and periods but do not attempt open-domain, natural-language interpretation at the scale of the full question space. Second, **retrieval-augmented generation (RAG)** grounds language generation in an external corpus so that outputs are traceable to source material rather than to opaque parametric memory (Lewis et al., 2020; Karpukhin et al., 2020); this is directly relevant to a domain in which every assertion should be attributable to a classical rule and a computed fact. Third, **dense-retrieval semantic routing** — using sentence embeddings (Reimers & Gurevych, 2019) and efficient vector search (Johnson et al., 2019) to map an utterance to an intent — provides an alternative to brittle keyword taxonomies and to costly per-query classifier calls with large models (Brown et al., 2020).

Relative to a monolithic large-language-model approach, an explicitly decomposed agent federation offers three advantages in this domain: (i) **auditability** — each answer can be traced to a specific agent, its declared primitives, and the rules it retrieved; (ii) **controllability** — a faulty interpretation can be corrected in one agent's specification without retraining a model; and (iii) **faithfulness** — grounding generation in retrieved classical rules and engine-computed facts constrains the well-known tendency of free-running models to fabricate. Our contribution is not a new learning algorithm but an architecture that composes established components (embedding retrieval, RAG, chain-structured reasoning; Wei et al., 2022) around a faithfully digitized classical corpus.

4. System Overview

Figure 1 shows the end-to-end pipeline. A user's free-text question enters a two-stage semantic router, which selects the responsible agent(s) from the federation of 1,000. The chosen agent declares which computational primitives it needs; the computational Jyotiṣa engine, having already calculated the sidereal chart from the user's birth data, supplies exactly those facts. The agent retrieves the applicable classical rules from a vectorized rule corpus, and a grounded synthesis layer composes the answer strictly from the computed facts and retrieved rules, subject to guardrails. Nothing in the answer originates outside the retrieved rules and computed facts.

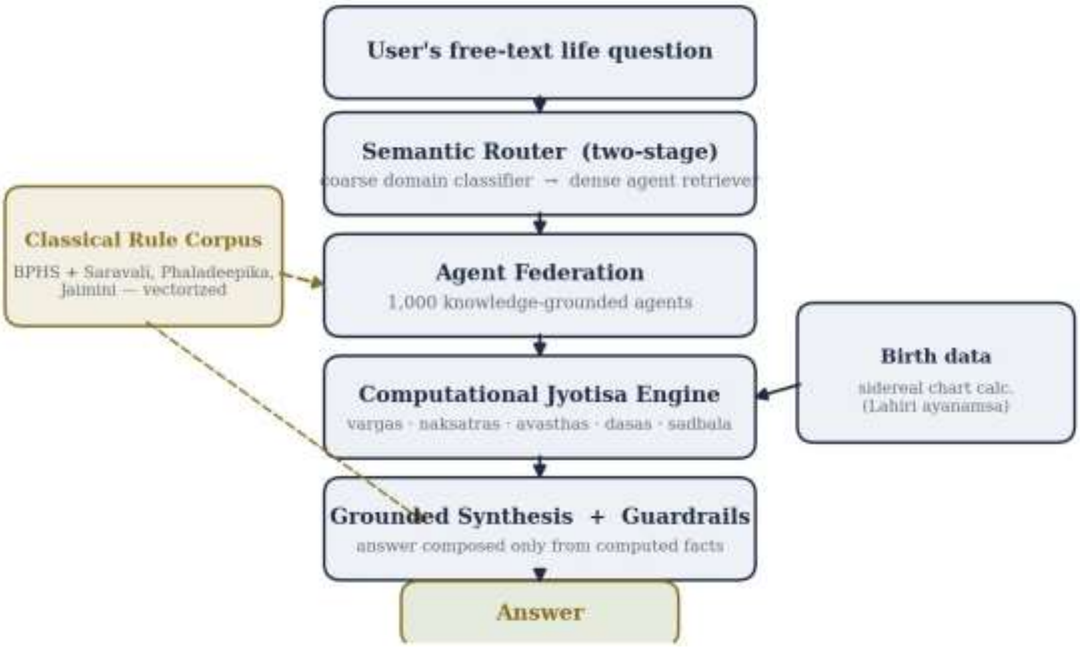

Figure 1. The RetroGrade knowledge-grounded question-answering pipeline. The classical rule corpus (left) grounds both agent selection and answer synthesis; the birth-data chart calculation (right) feeds the computational engine.

5. The Question Space: A 14,000-Item Life-Domain Taxonomy

The design begins not from the texts but from the *user*. We enumerated the questions a person is realistically likely to bring to a consultation and organized them into a three-level taxonomy: thirteen top-level **life domains**, each divided into **sub-domains**, each populated with concrete **canonical**

questions and their paraphrase families. The leaf level contains approximately 14,000 canonical questions; with paraphrase variants the routable surface is substantially larger. Table 5 gives the top-level distribution.

Life domain	Approx. canonical questions	Agents
Marriage & Relationships	1,850	130
Career & Profession	1,700	120
Wealth & Finance	1,550	110
Health & Longevity	1,400	100
Timing & Daśā Analysis	1,350	95
Education & Learning	1,100	80
Progeny & Family	980	70
Spirituality & Mokṣa	820	60
Personality & Character	780	55
Property, Vehicles & Assets	760	55
Foreign Travel & Relocation	640	45
Litigation & Adversity	620	45
Remedies & Upāya	450	35
Total	≈ 14,000	1,000

Table 5. Top-level life-domain distribution of the question taxonomy and the agent federation. Counts are approximate and reflect the relative density of real user demand observed during taxonomy construction.

Two design principles govern the taxonomy. First, **coverage over elegance**: it is better to over-enumerate near-duplicate questions than to leave a plausible user intent unrouted, because an unrouted question degrades the user experience far more than a redundant one. Second, **procedure-aligned partitioning**: sub-domains are drawn so that all questions within a sub-domain are answered by the *same* underlying analytical procedure. This second principle is what allows a clean one-or-few mapping from sub-domains to agents.

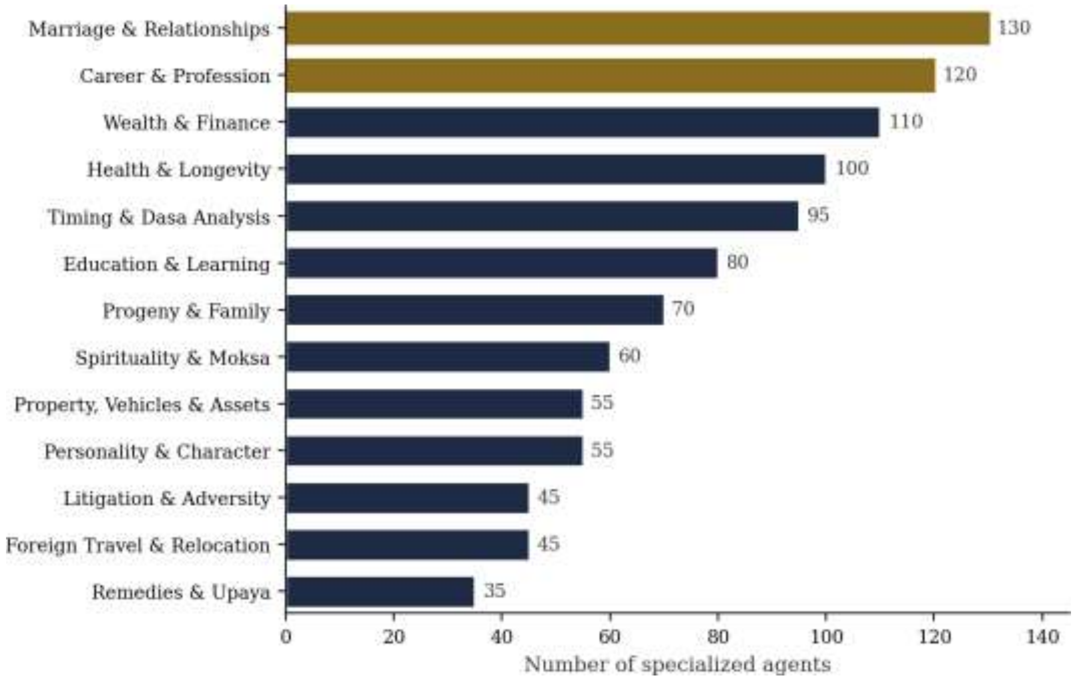

Figure 2. Distribution of the 1,000 agents across the thirteen life-domain clusters. Marriage/relationships and career are the densest clusters, mirroring observed user demand.

6. Agent Design: One Thousand Knowledge-Grounded Agents

An *agent* in our system is not an autonomous planner but a **scoped, declarative analytical procedure**: a specification that binds a family of user intents to the exact classical machinery that answers them. Each agent is defined by the schema in Table 6.

Field	Content
Scope	the sub-domain of intents the agent answers (e.g. "timing of first marriage")
Significators	the houses, house-lords and naisargika kārakas the question depends on (e.g. 7th house, its lord, Venus)
Authoritative vargas	the divisional charts required for corroboration (e.g. D-1 and D-9 for marriage)
Required primitives	which of nakṣatra, avasthā, ṣaḍbala and daśā systems the procedure consults
Rule references	pointers into the vectorized classical corpus (BPHS chapters and allied texts)
Yoga checks	named planetary combinations to test for (e.g. Rāja-yogas, Dhana-yogas, doṣas)
Output contract	the structure of the answer: claim, supporting configuration, strength/confidence, timing, caveat

Table 6. The agent-specification schema. Every agent declares its significators, authoritative divisional charts, required primitives and rule references, making its reasoning fully auditable.

Agents were derived from the taxonomy by **procedure clustering**: sub-domains that share significators, authoritative vargas and rule references collapse to a single agent, while a sub-domain whose questions branch into distinct procedures (for example, “will the marriage happen” versus “when will it happen” versus “what will the spouse be like”) expands into several. The result is 1,000 agents whose granularity tracks analytical distinctness rather than surface topic. Because each agent is a declaration rather than a trained model, agents are independently versioned, reviewed and corrected — an error in, say, the property-timing agent is fixed by editing one specification, with no effect on the other 999.

Worked example

Consider the question “When will I get married?” It routes to the *first-marriage-timing* agent, whose specification names the 7th house, its lord and Venus as significators; D-1 and D-9 as authoritative; Viṁśottarī and Yoginī daśās plus the 7th-lord's ṣaḍbala and avasthā as required primitives; and the relevant BPHS rules on marriage timing as references. The engine returns those exact facts for the user's chart; the agent tests the named yogas and the daśā windows in which the 7th-lord and Venus are activated and sufficiently strong; and the synthesis layer reports the most probable window, the configuration supporting it, the cross-daśā agreement as a confidence signal, and an explicit caveat where the chart is ambiguous.

7. Semantic Routing

The router's task is to map an arbitrary free-text question to the agent(s) that can answer it. An earlier design used a keyword taxonomy backed by a large-model classifier call on every query; this was brittle to paraphrase and slang, and imposed a latency and cost penalty on each interaction. The current design replaces both with dense-embedding retrieval.

At build time, every canonical question and paraphrase in the taxonomy is embedded with a sentence encoder (Reimers & Gurevych, 2019) and indexed for approximate nearest-neighbour search (Johnson et al., 2019), with each vector carrying its agent label. At query time the user's question is embedded once and the index is searched, yielding a ranked list of candidate agents by semantic proximity. Routing proceeds in two stages: a **coarse stage** localizes the query to one or a few life-domain clusters, pruning the search space; a **fine stage** resolves the specific agent within those clusters. A lightweight lexical signal (BM25; Robertson & Zaragoza, 2009) is fused with the dense score to protect rare but decisive domain terms — proper nouns of yogas, nakṣatras or texts — that pure embeddings can under-weight.

Three behaviours make the router robust in production. **Multi-agent fan-out:** a genuinely compound question (“how are my career *and* health this year?”) is dispatched to more than one agent and the results merged. **Abstention:** when the top candidates fall below a similarity threshold, the router declines to guess and asks a clarifying question rather than mis-routing. **Continuous enrichment:** unmatched or low-confidence questions are logged and, once reviewed, folded back into the taxonomy as new paraphrases or, where warranted, new agents — so coverage grows monotonically with usage. Eliminating the per-query classifier call also removed a redundant model invocation from the hot path, reducing both latency and cost without loss of routing quality.

8. The Computational Jyotiṣa Engine

The engine is the deterministic substrate beneath the agents. From the user's date, time and place of birth it computes the sidereal chart using established ephemeris routines with the Lahiri (Chitrapakṣa) ayanāṁśa, the standard reference in modern Indian practice. It then derives, on demand, every primitive an agent may request:

- **All sixteen vargas** (Table 1), so that any agent's corroboration requirement can be satisfied;
- **Nakṣatra and pāda** positions for every graha and the lagna, at the resolution personality and compatibility agents require;
- **Avasthā states** under the Bālādi, Jāgradādi, Dīptādi and Lajjitādi schemes (Table 2), as modifiers on every rule;
- **Daśā timelines** — Viṁśottarī, Yoginī, Aṣṭottarī, Kālachakra and Chara — nested to the depth a timing question needs (Table 3);
- **Full ṣaḍbala** with all six components (Table 4), including **vakrī status and ceṣṭā-bala** for the five true planets;
- **Aspects (dṛṣṭi)**, planetary friendships, combustion (astaṅgata), planetary war (graha-yuddha), and the standard library of named **yogas** and **doṣas**.

Determinism here is a virtue: because the engine's outputs are exact functions of the birth data, every interpretive statement can be traced back to a reproducible numerical fact. The agents supply judgement; the engine supplies ground truth.

9. Grounded Synthesis and Guardrails

The final stage composes a natural-language answer. It receives the agent's structured findings — the configuration, the strengths, the daśā windows, the fired rules and their classical references — and renders them into readable prose that follows the agent's output contract (claim, support, strength, timing, caveat). Two constraints govern this stage. **Grounding:** the synthesizer may assert only what the computed facts and retrieved rules support; it is instructed and checked not to introduce claims that do not trace to that evidence. **Register and safety:** because the domain touches health, mortality, relationships and money, the synthesizer is bound by guardrails — it avoids deterministic or alarming verdicts on medical and mortality questions, frames sensitive findings as tendencies rather than certainties, presents classical disagreements (as in the retrograde case of Section 2.6) transparently, and defers to qualified professionals on legal, medical and financial specifics. These guardrails are a design requirement of a consumer advisory system, independent of any position on the tradition's validity.

10. Implementation

The taxonomy, agent specifications and rule references are maintained as versioned, declarative artifacts, so that domain review and correction are decoupled from software deployment. The routing index is built offline from the embedded taxonomy and served through an approximate nearest-neighbour store. Chart computation is performed by the deterministic engine described in Section 8. Answer synthesis is performed by a large language model constrained to the grounded, guarded contract of Section 9; retrieval-augmented prompting supplies the model with the fired rules and computed facts at generation time so that its output remains attributable. The separation of concerns — declarative knowledge, deterministic computation, retrieval, and constrained generation — is deliberate: it localizes every class of error to a single, independently testable layer.

11. Evaluation

We distinguish sharply between what *can* be evaluated objectively and what cannot. The **predictive truth** of astrological statements is outside the scope of this paper and, we would argue, outside the reach of the system's own design goals. What *is* objectively measurable, and what a system of this kind must be held to, is: (i) whether questions are routed to the correct agent; (ii) whether answers are faithful to the classical rules and computed facts they cite; and (iii) whether coverage of the question space is adequate. We propose the protocol below and report indicative development-set figures; authors preparing a peer-reviewed submission should replace the indicative values with measurements on a frozen production build and a held-out, human-annotated test set.

11.1 Routing accuracy

On a held-out set of human-labelled questions (including deliberate paraphrases, code-mixed Hindi-English utterances and slang), we measure top-1 and top-3 routing accuracy against the gold agent label, and the abstention rate on out-of-scope questions.

Metric	Definition	Indicative (dev)
Top-1 routing accuracy	fraction routed to the correct agent at rank 1	[to be measured]
Top-3 routing accuracy	correct agent within the top three candidates	[to be measured]
Paraphrase robustness	top-1 accuracy on held-out paraphrases	[to be measured]
Out-of-scope abstention	fraction of out-of-scope questions correctly declined	[to be measured]

Table 7. Routing-accuracy evaluation protocol. The indicative column is intentionally left for measured values; the taxonomy's gold labels make this an ordinary supervised evaluation.

11.2 Rule-faithfulness

For a stratified sample of (question, chart) pairs, a qualified Jyotiṣa reviewer judges whether every assertion in the answer is supported by the cited rule and the computed configuration, and whether any required corroboration (e.g. the authoritative varga) was actually performed. We report the proportion of assertions rated *supported*, *unsupported* or *contradicted*, and the proportion of answers with no unsupported assertions. This directly measures the grounding constraint of Section 9 and is, in our view, the single most important quality metric for a system of this kind.

11.3 Coverage

Coverage is estimated by sampling real user questions and measuring the fraction that route above the confidence threshold without abstention, tracked over time as the continuous-enrichment loop (Section 7) folds unmatched questions back into the taxonomy. A healthy system shows coverage rising monotonically and abstention concentrating on genuinely out-of-scope requests rather than in-scope paraphrases.

12. Discussion

Decomposing the domain into an enumerated taxonomy and a federation of declarative agents trades the apparent generality of a single large model for **auditability, controllability and faithfulness**. In a sensitive advisory setting these properties matter more than raw generality: a wrong but confident answer from an opaque model is worse than a narrower answer whose every step can be inspected and corrected. The architecture also makes the classical tradition itself more legible — because each agent must declare the significators, vargas and rules it uses, the system doubles as an executable, testable encoding of BPHS-based practice, in which disagreements among authorities (retrogression being the paradigm case) are represented explicitly rather than smoothed over.

The approach generalizes beyond Jyotiṣa to any large, rule-based interpretive tradition with computable primitives and an enumerable question space — other astrological systems, classical medical or legal corpora, or scripture-based advisory practices — wherever faithfulness to a source tradition and auditability of each answer are prized over unconstrained generation.

13. Epistemological Positioning, Ethics and Limitations

We state our position without equivocation. **This paper makes no claim that astrological predictions are empirically valid.** Astrology is not accepted as a predictive science by the scientific mainstream, and nothing in our architecture constitutes evidence for or against the tradition's claims about the world. The scientific object of this work is the **faithful computational modeling of a classical interpretive system** and the natural-language-processing architecture that renders it usable at scale. Our objective evaluation targets — routing accuracy, rule-faithfulness, coverage — deliberately concern fidelity to the tradition and to the user's intent, not correspondence between predictions and outcomes.

This positioning carries ethical obligations that are built into the system rather than bolted on. The guardrails of Section 9 exist because users may bring questions about illness, death, and financial or legal crises; on such questions the system must avoid deterministic or alarming verdicts, frame findings as tendencies, and direct users to qualified medical, legal and financial professionals. The platform presents itself as a tool for reflection grounded in a cultural tradition, not as a substitute for professional advice.

The principal limitations follow from the design. Coverage, while broad, is bounded by the taxonomy and grows only as the enrichment loop incorporates new intents. Interpretive quality is bounded by the fidelity of the encoded rules and by the classical disagreements the tradition itself contains; where the authorities differ, the system surfaces the divergence but cannot resolve it. Finally, the objective metrics we can report speak to fidelity and routing, not to real-world predictive accuracy, which we make no attempt to assess.

14. Conclusion and Future Work

We have described a knowledge-grounded, multi-agent architecture that operationalizes Parāśarī Jyotiṣa for open-ended natural-language question answering. By decomposing the domain into a ~14,000-item life-question taxonomy and a federation of 1,000 declarative agents — each grounded in Bṛhat Parāśara Horā Śāstra and each invoking the appropriate divisional charts, nakṣatras, avasthās, daśā systems and ṣaḍbala components, with retrograde motion and ceṣṭā-bala as a first-class signal — the system achieves broad coverage while keeping every answer auditable and correctable. Future work includes reporting the full evaluation of Section 11 on a frozen production build, extending the enrichment loop with active-learning selection of the highest-value unmatched questions, and formalizing the agent-specification schema as an open interlingua for encoding classical interpretive rules.

15. Acknowledgements

The authors thank the practitioners and reviewers who contributed to the construction and validation of the question taxonomy and agent specifications.

References

1. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., … Amodei, D. (2020).

Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.
2. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186.
3. Jaimini. (n.d.). Jaimini Sūtras (S. Rath, Trans.). Classical Sanskrit treatise on the Jaimini system of Jyotiṣa.
4. Johnson, J., Douze, M., & Jégou, H. (2019). Billion-scale similarity search with GPUs. IEEE Transactions on Big Data, 7(3), 535–547.
5. Kalyāṇavarma. (n.d.). Sārāvalī (R. Santhanam, Trans.). Ranjan Publications. (Classical treatise, c. 8th century CE.)
6. Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W. (2020). Dense passage retrieval for open-domain question answering. Proceedings of EMNLP 2020, 6769–6781.
7. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., … Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 33, 9459–9474.
8. Mantreśvara. (n.d.). Phaladeepikā (G. S. Kapoor, Trans.). Ranjan Publications. (Classical treatise on predictive astrology.)
9. Parāśara. (n.d.). Bṛhat Parāśara Horā Śāstra (R. Santhanam, Trans., Vols. 1–2). Ranjan Publications. (Foundational treatise of Parāśarī Jyotiṣa.)
10. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks.
11. Proceedings of EMNLP-IJCNLP 2019, 3982–3992.
12. Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. Foundations and Trends in Information Retrieval, 3(4), 333–389.
13. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017).
14. Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.
15. Varāhamihira. (n.d.). Bṛhat Jātaka (Classical treatise on natal astrology, c. 6th century CE).
16. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35, 24824–24837.