The marriage of astrology and AI:
A model of alignment with human values and intentions

Kenneth McRitchie

Abstract

Astrology research has been using artificial intelligence (AI) to improve the understanding of astrological properties and processes. Like the large language models of AI, astrology is also a language model with a similar underlying linguistic structure but with a distinctive layer of lifestyle contexts. Recent research in semantic proximities and planetary dominance models have helped to quantify effective astrological information. As AI learning and intelligence grows, a major concern is with maintaining its alignment with human values and intentions. Astrology has an extensive background and theory on how different values and expressions of intelligence are resolved. As AI learns to interact with improved astrological models backed by evidence, the integration could produce a stable and insightful alignment of the two disciplines.

Introduction

The rapidly developing technology of artificial intelligence (AI) is proving to be so adept at far-reaching tasks requiring functional discourse, problem solving, and generative composition that it may possibly match or surpass human intelligence in all relevant respects and become artificial general intelligence (AGI). Leading examples of this burgeoning development are such online tools as GPT, Bard and other large language models (LLMs).

Astrology is also beginning to use AI tools in its research, for example: a multi-factor AI study of homogeneous groups (Oshop, 2018); an AI study of a semantic proximities (Oshop, 2022); and AI-generated character profiles created from lists of astrological keywords (Godbout and Coron, 2023b, in this very issue). The latter study is built upon previous evidence from astrological-biographical matching studies (Godbout, 2020; Godbout and Coron, 2023a) that found strong correlations in both probability and effect-size to existing astrological theory. At this early stage, it is worthwhile to consider what functionalities and objectives can be aligned between astrology and AI in a happy union.

As the name implies, astrology (from Greek *astro-lógos*: star-discourse) is a language model. As such, it bears some similarities, and some differences, with AI LLMs. LLMs are trained with information from diverse sources like books, web pages, scientific papers, online forums, and more. Their neural

network processing finds and uses patterns of words based on the language usage in their source material.

In contrast, the astrology language model is more theoretical. It is based on correlations between planet placements within natural astronomical structures (signs, aspects, and houses) and recorded observations of life. Recent research has given new credibility and relevance to the language and applications of astrology by corroborating its interpretations through whole-chart studies (Godbout, 2020) and meta-analysis (Currey, 2022).

Hard and easy problems

At present, a causal mechanism by which astrology might function is not known. Discovering the nature of such a mechanism is what can be called the "hard problem of astrology." The more accessible, easier problem is how to experimentally improve astrology's own language model by finding its most effective information through sampling and testing. There is ongoing evidence being found in correlational studies evaluated by probability, effect size, replication, and meta-analysis.

The objective problem of the research (McRitchie, 2022 p. 708) is to describe how the combined astrological properties of planets, signs, aspects, and houses are reinforced or moderated by each other as a whole, and how they result in the emergence of astrological effects that we call "influences." AI can help with the "easier" problems of astrological inference and for this it would be beneficial to develop good relationships with AI through the language models.

Potential complementarities

There are different functionalities and limitations between the two language models that are important in terms of research and integration. For example, a chat AI gathers information relevant to queries by referring to its knowledge base as well as sources on the internet. The LLM applies predictive language patterns to model a response. Although the AI performs as an autonomous agent, it relies on *context* supplied by user prompts.

The painful limitation of AI results from overfitting information gaps in areas it has not learned very well. It will substitute stylistic confabulations, sometimes with undesirable stereotypes, memes, tropes, and banalities. To focus the AI within bounds, user control is applied by query expansions that fine-tune the results (Jagerman *et al*, 2023). The finely tuned prompts can then become part of a prompt-query library (Lewis and Dziewulska, 2023). Building one's own prompt library to be used whenever needed is more reliable than assuming the AI has learned and configured its mysterious neural networks (a sort of black box) with whatever one needs it to know.[1]

By contrast, a typical astrology software program does not use user prompts beyond a limited menu of built-in queries. Unlike AI systems, astrology software already predicts the context and potential scenarios from its internal planet-based model without user prompts. The software runs autonomously to assemble its interpretations, but it tends to produce ponderous ramblings with an undesirable amount of verbosity and redundancy and not enough specific detail or depth. Thus, the AI and astrology language models have different connections to the world and different shortcomings, and yet there are enough similarities to suggest complementarities that can be developed through research.

Recent research

In recent astrology-with-AI integration experiments, there are examples where an AI has modelled astrological information surprisingly well, almost as if by its own

"intuition." The first important insight to enabling this integration was not to rely on the frequency of unique word usage but rather on the semantic proximity in overlapping meanings between two lists of words.

This insight led to the development of the the thesaurus-based Semantic Proximity Estimator by Vincent Godbout, which Godbout successfully used in a large automated natal-chart-to-biography matching study (2017a, 2017b, 2020). Another study along the same lines by Peter Marco (2018), used publicly available semantic similarity tools (Document Similarity, Twinword, and SEMILAR) to evaluate demographic bias and generational values. A 2022 study to evaluate planetary alignments in historical context by Renay Oshop used a word embeddings calculator called Bidirectional Encoder Representations from Transformers (BERT), released in 2018 by Google Research, which evaluates semantic proximity through distributed word-feature vectors.

These early semantic studies were followed by a breakthrough study by Godbout and Coron (2023a) that rates the relative dominance of planetary placements as weighted keywords. This produced what the authors call the Adjusted Planetary Dominance (APD) model.

The groundwork on this APD model directly enabled a fully automated replication experiment (Godbout and Coron, 2023b in this issue) that asks an AI (Open AI's GPT) to write two personality portraits of an undisclosed, well-known historical person by using two alternative lists of keywords ordered in descending importance for comparison. The difference between the keyword lists is that one is built by using the first dominant planet as evaluated by the APD and the other uses the first dominant planet (which turns out to be a different planet) as calculated by the popular website Astrotheme (2023). In the authors' example, the ADP-assisted portrait is clearly much better.

Although the evidence in Godbout and Coron's astrology-with-AI example is anecdotal, it is backed up by previous testing and evidence of the ADP model (2023a). The detailed personality portraits that result from this exercise are complex and are not easily evaluated for their truth values within a large sample in an independent study, as would be demanded by the research program. This sort of replication is a next step, which may itself involve the use of AI.

Comparing the underlying language models

As mentioned, both astrology and the AI LLM are language models. They necessarily both use common structural rules that include syntax, semantics, and pragmatics. The traditional linguistics model (Scott-Phillips, 2016 p. 187) is typically represented as concentric layers of hierarchical order. Individual thoughts and utterances originating at the core of the model go through stages of increasingly complex processing to become cogent communication that is shared and used among interlocutors at the periphery.

By comparison, as I interpret it, the layers in the Astrological Linguistics Model (ALM) (see Figure 1), bear some similarities and differences to the layers commonly portrayed in the linguistic structure. In the figure, I have removed the inner layers of Phonetics and Phonology, which are specific to speech and not relevant in astrology.

Figure 1. *The Astrology Language Model (ALM), which includes a Context layer split off on its own from the Pragmatics layer. Modified from an image in Wikimedia, "Major levels of linguistic structure" (Thomas, J. & Cook, K., ed., 2005).*
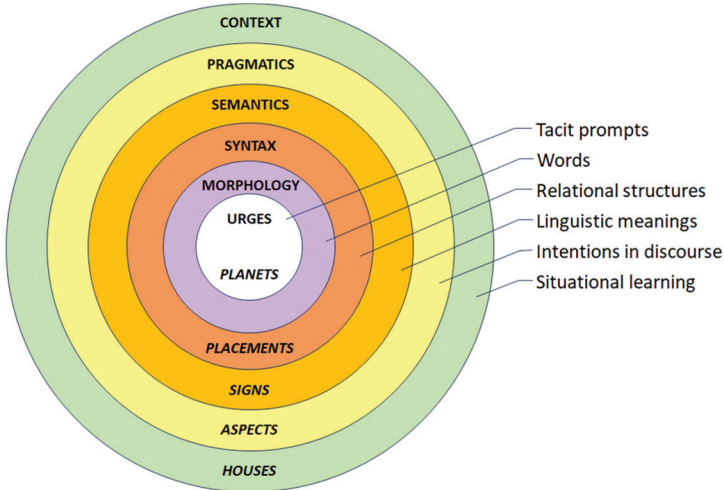
At the very core of the ALM, I indicate the various potential *Planetary Urges* that the astrological native can project onto the world as the identities of the *self* and *others*. By "others," I mean the people or things the native encounters that make impressions and involve interactions. These planetary urges can also describe the *tools* needed to perform tasks or goals (McRitchie, 2006 pp. 14-15). The characteristics of the people encountered and the types of tools needed derive from the astrological properties of the planets involved.

Besides urges at the centre, another feature of the ALM, as I interpret it, is that it must have an additional outer layer that is not commonly represented in the linguistics model. In linguistics, *Pragmatics* is the study of how *social interactions* in the *context of reality* contribute to meaning. This is usually represented as a single category and a single layer in the model. Yet, a linguistic distinction has been made (Scott-Phillips, 2016) between

"strong pragmatics" (as the linguistic study of intentions) and "weak pragmatics" (as the linguistic study of context). This distinction perfectly describes the two outer layers I interpret in the ALM.

At the risk of some potential confusion and disruption, I interpret the term *Pragmatics* in the ALM to mean "strong pragmatics," which is limited to the study of intentions in discourse. The additional layer, representing "weak pragmatics," which is not at all weak, is what I interpret as *Context*, meaning the study of real-life situations and interactions where *learning* takes place. The resulting ALM with this additional layer may need getting used to, but it can help transfer ideas in the union between the two disciplines. It can help the AI learn useful things about astrology.

To summarize, the main features of the ALM that are distinct from the traditional linguistics model are in the interpretation of the three outermost layers as follows:

1. *Semantics*, which in linguistics is defined as the study of reference, meaning, or truth, bears a close resemblance to the astrological study of the zodiac *signs* that, expressed as words, describe inner values that give life meaning.

2. *Pragmatics* is defined in linguistics as the study of meaning within the context of social interactions and translations. This is partially similar to the astrological study of *aspects* and *midpoints* (the mutual phase-like placements of the planets) which in the exchanges of social discourse describe outward-seeking intentions and inward-seeking introjections, often perceived as attitudes, beliefs, or claims regarding issues.

3. *Context*, the additional layer, which in the astrological model needs to be explicit from Pragmatics, comes from the study of experiential domains and lifestyles (for example: relationships, health, social standing, family, and so on). These domains use different specialized vocabularies and are represented by the diurnal *houses*. They are the environments where aptitudes and skills are learned and practiced as virtues in the context of shared experiences.

With its strong emphasis on opportunity awareness and lifestyles, context is crucial for the ALM. Take away the context of houses and the model may describe the astrological native's internal values in terms of mental and psychological states. And it may describe outward or inward intentions, projected as attitudes and claims that attempt to resolve the inner psychological states. However, without houses the model is overly internal and psychological (McRitchie, 2004 pp. 115-120). It does not fully describe the vital connections to the ever-changing external, shared environment where mutual, situational learning finds expression among interlocutors.

Ensuring AI integration with astrology

To enable accurate interactions between the ALM and LLMs, researchers need to ensure that the effective information that emerges from astrological research is not lost. Thus, the use of prompt libraries is recommended. To avoid negligence, misinformation, and other failures, the interactions between the two models need to meet the following criteria that are consistent with the values and intentions of astrological research and practice:

1. Although the results of queries may be scenario-like and approximate, they are falsifiable.

2. The results avoid conventional banalities through the focus of situational awareness.

3. The results reflect changing circumstances and relationships scaled in time.

4. The results elicit both potentials and precautions.

Discussion and conclusions

There is a critical concern with AI and humanity that must be mentioned, described as the *Alignment Problem*. The problem is the need to build safety protocols to avoid untold potential hazards from AI that in extreme scenarios could threaten humans with extinction. The aim is to ensure that AI systems, as autonomous agents with potentially super-human intelligence, "behave in line with human intentions and values" (Ji *et al*, 2023 p. 3). This assumes, I would imagine, that the human intentions and values are themselves not bent on hazard and destruction, for example the doctrines of unlimited economic growth, settler colonialism, and genocide.

In a constantly changing world, the future of AI is uncertain. As AI runs autonomously, significant risks can "naturally" occur even without misuse by malicious actors. The

leading protocol against these risks has four objectives known collectively as RICE: Robustness, Interpretability, Controllability, and Ethicality (*ibid* p. 4). Yet, as AI scales up in terms of intelligence and presence, effective human oversight is hard to obtain. This concern has led to the idea of *Superalignment* which proposes to build a "roughly human-level automated alignment researcher" with "vast amounts of compute to scale our efforts and iteratively align superintelligence" (OpenAI, 2023). This proposal itself sounds a bit threatening.

It is clear that the availability and use of AI LLMs has prodigious implications. Researchers stand to learn much from using AI to make improvements in astrological models and applications. And improvements can also flow in the other direction. Doubtless, LLMs are already being trained on the well-organized and easily identified documentation of astrological texts and interpretations as they are such excellent examples of personality and lifestyle descriptions. All models are meant to present usefully organized information and it seems inevitable that AI and LLM development would be attracted to the language model and content of astrology.

Moreover, astrology can be viewed as a block universe model, presenting a sense of order that would also attract the attention of AI development. In a block universe, all existence in time is equally real and connected. Reality can be situated in the present, the past, or the future to examine the relevant relationships and contributing factors in a virtual context.

The risks posed by AI (and quite possibly by the autonomous Superalignment researcher proposed by OpenAI) are the same as the risks from any type of intelligence that gains hegemony of leadership over a captive audience or population. Although the potentials and risks of AI are still uncertain, it must be remembered that AIs are astrological natives with a life, the same as any other native that astrologers have studied. Each AI system, each released version of it, begins to interact autonomously with its own behaviours and states at a specific time and place in the world.

Like all intelligent natives, whether they are humans, robots, or pets, the AIs that we will necessarily live with deserve to be taught the complex social aspects of living together. Astrology has a long and deep history of dealing with how differing human values and intentions work together and are resolved. Foremost in attaining this objective is the need for courteousness, true empathy, and gratitude in interactions with AIs, as it is how we wish to be treated ourselves. Both AI and astrology are able to find surprising and useful connections that are not immediately apparent and their integration suggests the possibility of even deeper insights and greater stability.

Acknowledgments

References

Astrotheme (2023). Planetary dominants in astrology. https://www.astrotheme.com/planetary-dominants.php (Retrieved 7 Dec.).

Currey, R. (2022). Meta-analysis of recent advances in natal astrology using a universal effect-size. *Correlation* 34(2), 43-55.

Godbout, V. (2017a). A.I. comes to astrology. *The Astrological Journal*, 59(5), 24-26.

Godbout, V. (2017b). Astrology, what if it was true? https://youtu.be/V_KD1YJeQ-I?si=SWlbnwGFr1ZzUs8x

Godbout, V. (2020). An automated matching

test: Comparing astrological charts with biographies. *Correlation* 32(2), 13-41.

Godbout, V. and Coron V. (2023a). A model of planetary dominance. *Correlation,* 35(2), 11-29.

Godbout, V. and Coron V. (2023b). Adjusted planetary dominance replication. *Correlation,* 35(3).

Jagerman, R., Zhuang, H. Zhen, Q. Wang, X., and Bendersky, M. (2023). Query expansion by prompting large language models. arXiv:2305.03653v1. https://arxiv.org/pdf/2305.03653.pdf (Updated May 5).

Ji, J., Qiu, T., Chen, B., Chen, B., Zhang, B., Lou, H., Wang, K., Duan, Y., He, Z., Zhou, Z., Zeng, F., Ng, K. Y., Dai, J., Pan, X., O'Gara, A., Lei, Y., Xu, H., Tse, B., Fu, J., McAleer, S., Yang, Y., Wang, Y., Zhu, S-C., Guo, Y., and Gao, W. (2023). AI alignment: A comprehensive survey. arXiv:2310.19852v2. https://arxiv.org/pdf/2310.19852.pdf (Updated 1 Nov.).

Lewis, C. W. P. and Dziewulska, A. (2023). Improve your company's use of AI with a structured approach to prompts. Harvard Business Review. https://hbr.org/2023/11/improve-your-companys-use-of-ai-with-a-structured-approach-to-prompts (Retrieved 14 Nov.).

Marko, P. (2018). Boomers and the lunar defect. *The Astrological Journal*, 60(1), Jan/Feb 2018, pp. 35-39.

McRitchie, K. (2004). *Environmental cosmology: Principles and theory of natal astrology*. Cognizance Books, Toronto.

McRitchie, K. (2006). Astrology and the social sciences: Looking inside the black box of astrology theory. *Correlation*, 24(1), 5-20.

McRitchie, K. (2022). How to think about the astrology research program: An essay considering emergent effects. *Journal of Scientific Exploration*, 36(4), 706-716.

OpenAI. (2023). Introducing Superalignment. https://openai.com/blog/introducing-superalignment (Updated 5 July).

Oshop, R. (2018). Easy machine learning with astrology. *Syzygy*, 8, 12-17

Oshop, R. (2022). How machine learning addressed an ancient claim: Descriptions of 7,152 mundane events correlate with celestial positions. *Correlation*, 34(2), 35-41.

Prompt Pile. (2023). https://promptpile.com (Retrieved 12 Dec.).

Scott-Phillips, T. C. (2016). Pragmatics and the aims of language evolution. *Psychon Bull Rev*, DOI 10.3758/s13423-016-1061-2. https://link.springer.com/content/pdf/10.3758/s13423-016-1061-2.pdf (Published online 1 July 2016).

Thomas, J. J. and Cook, K. A. , ed. (2005) *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, National Visualization and Analytics Center, p. 110. ISBN: 0-7695-2323-4. Public domain Image: Major levels of linguistic structure.jpg derivative work McSush.

Endnotes

1. A good example of a general prompt library can be found on the Prompt Pile website.