Inferring Social Structures From Language

Keywords: Social Structuralism, Knowledge Graphs, Causal Reasoning, Linguistic Structuralism, Explanations

Extended Abstract

A central idea of structuralism is that a thing's identity is not defined by its intrinsic properties, but by the larger structure it is part of. This is true if we want to understand human behavior in the context of larger social structures, but it is also true for the conceptual structure of language, in which concepts have relations with each other. Further, the meaning of behaviors, concepts and words can be derived from its place in the structure. To quote Sally Haslinger "*sometimes it is good and useful to explain the behavior of a thing by explaining the behavior of something of which it is part, if it is a part of whose behavior is constrained by other parts of the whole*". Applying this principle to language we developed analytical methods and technological capabilities that allow us to parse unstructured text into relational graphs. The graphs allow us to draw long-range connections beyond the language into knowledge about the social structures. Drawing on theories from linguistic anthropology, sociology, philosophy (Haslanger, 2015) and explanatory pluralism (Shtulman & Lombrozo, 2016) we seek to provide theoretically informed, but computationally derived explanations for social structures from language. We call the system that integrates social theory into deep learning WEBER (see Figure 1). Next, we explain our NLP architecture and then provide an example using written explanations for behavioral health choices from mothers in India. We show how textual explanations can be used to understand underlying beliefs and social structures using NLP.

For example, in the sentence "Anita and her bay live in India", "live in" is considered a 'social act' of which Anita is the agent of, and India the location where this act happens. It is inferred that Anita owns her baby (see Figure 1). In this simple example, the position of a word (e.g., representing a person, social actor, or identity) is part of a larger set or words (e.g., social institutions like nation states or kinship; more detail shown in Figures 3 and 4) that make up its situated meaning. Using the transformer-based SpEAR NLP architecture (Friedman et al., 2022) our system can encode text as vector *representations* and describe entities (e.g., social actors or social institutions) as nodes in a knowledge graph. Using theory, the WEBER system can *describe* the types of nodes and assign multiple attributes via multi-class labeling (Fellbaum, 2010). Next, we use social and cognitive theories to describe different types of relationships that define the relationship between those nodes. Specifically, our system is able to automatically provide diverse causal relations including qualitative increase/decrease ("Having more children *will increase* our expenses"), intentions and goals ("She is feeding the baby colostrum *to avoid* having diseases"), explicit rationales ("She fell ill *because* of contaminate drinking water"), temporal precedence ("Anita felt better *after* taking some medicine") or social influence relations ("Anita's friend must have *told* her to not drink the dirty water"). In sum, each knowledge graph contains *factors* (nodes or multi-node subgraphs) and *relations* (directed, labeled edges between factors). Each node of a factor may have multiple *attributes* that indicate multi-class labels and *word senses* estimating the WordNet synonym sets for that node, within the sentence context. Together factors (e.g., concepts, persons, institutions), attributes (describing those factors) and relationships between them provide us with a causal reasoning structure. The WEBER system can then extract patterns of causal explanations for behavioral decisions, social practices, positions, or perceptions.

For this study, we used a dataset of 10,000 responses explaining behavioral choices within the maternal health domain. The dataset used includes responses from interviews with mothers and Accredited Social Health Activists (ASHAs) in Bihar, India (Legare et al., 2020). Healthcare workers were prompted to provide explanations for behavioral choices by young mothers related to breast feeding, delivery, family planning and vaccinations. We identified 9 categories of causal factors (labeled as node/subgraph factor) and 6 types of causal relationships over a total of 8 different topics (Figures 2). The results of this analysis show differences for explanations for several behavioral choices. For example, mentioning actual diseases, strength/health, and well-being for characters as rationale for decisions that are aligned with the recommendations for healthcare workers (blue) versus decisions by young mothers not consistent with healthcare worker recommendations (orange). The welfare+ causal relations occur significantly more frequently in rationales consistent with healthcare worker recommendations, especially as it relates to iron supplements (IFA) and infant vaccination.

On the other hand, a characters' beliefs, fears, and social influence only appeared significantly as rationale for decisions that were opposed by healthcare workers (e.g., feeding cowmilk instead of colostrum). Fear is mentioned as a significant rationale for avoiding iron supplements, hospital births, family planning and vaccines. Likewise, a characters' *Beliefs* is more often cause of health decisions not consistent with healthcare worker recommendations. Finally, direct social influence *Social* only contributed significantly as an influence to deliver at home.

Overall, we find that health decisions consistent with health recommendations are explained with more objective rationales and health benefits, and decisions that go against health recommendations are more influenced by beliefs and fears. While this is but one specific example of how ML can be applied to understand factors influencing behavioral health decisions, the analysis and its outcomes have larger implications for how we can use language itself to understand the link between beliefs and larger social structures shaping behavioral choices.

References

Shtulman, A., & Lombrozo, T. (2016). Bundles of contra- diction: A coexistence view of conceptual change. *Core knowledge and conceptual change*, 49–67.

Legare, C. H., Akhauri, S., Chaudhuri, I., Hashmi, F. A., Johnson, T., Little, E. E., . . . others (2020). Perinatal risk and the cultural ecology of health in bihar, india. *Philo- sophical Transactions of the Royal Society B, 375*(1805), 20190433

Friedman, S., Magnusson, I., Sarathy, V., & Schmer- Galunder, S. (2022). From unstructured text to causal knowledge graphs: A transformer-based approach. *arXiv preprint arXiv:2202.11768*.

Haslanger, S. (2016). What is a (social) structural explanation?. *Philosophical Studies, 173*, 113-130.

Fellbaum, C. (2010). Wordnet. In *Theory and applications of ontology: computer applications* (pp. 231–243). Springer.

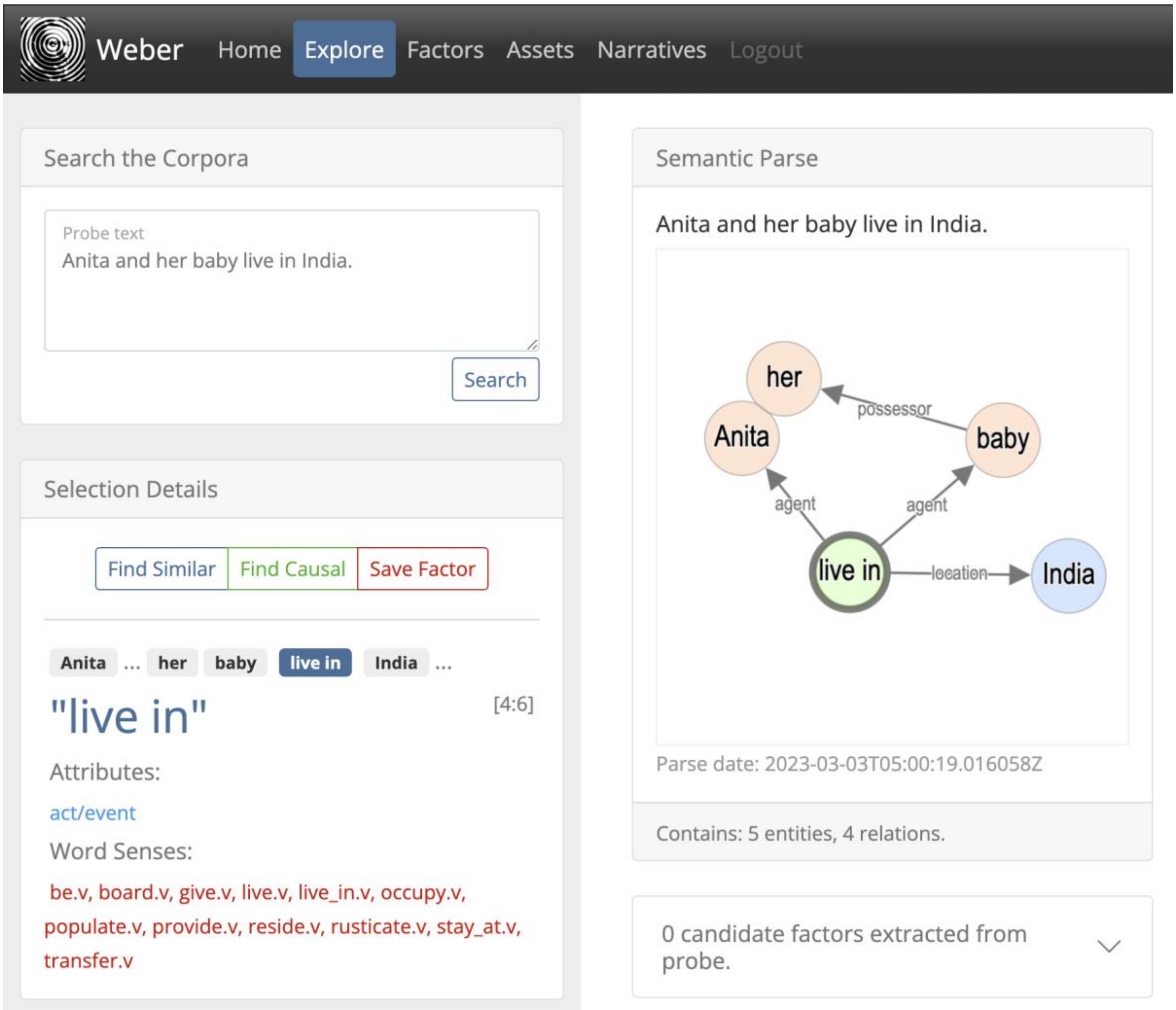

Figure 1. WEBER interface screenshot. WEBER automatically 1) parses sentences into contextually-informed factors, 2) ascribes relevant attributes (defining the concept that is described) and 3) extracts existing relationships and 4) is able to provide inferred relationships to larger social meanings (not pictured).

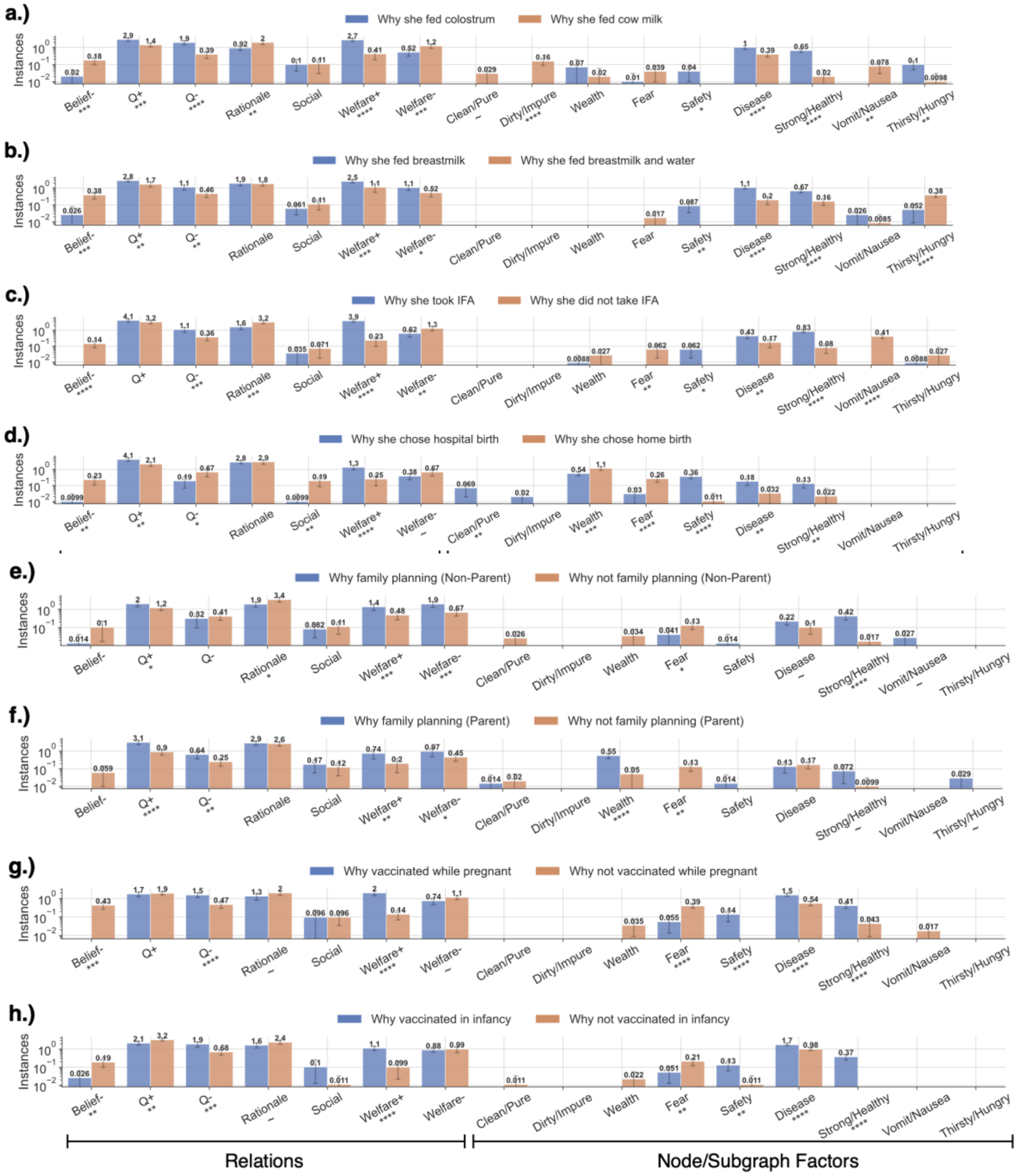

Figure 2. Results of explanations for behavioral health choices in 8 domains: (a) a mother chooses to feed her child colostrum *vs.* cow milk; (b) a mother chooses to feed her child breastmilk *vs.* breastmilk with water; (c) a pregnant woman chooses to take IFA tablets *vs.* not taking IFA; (d) a woman chooses to give birth at the hospital *vs.* at home; (e) a woman without children utilizes family planning *vs.* not; (f) a mother utilizes family planing *vs.* not; (g) a woman is vaccinated during pregnancy *vs.* not; (h) a woman vaccinates her infant *vs.* not.

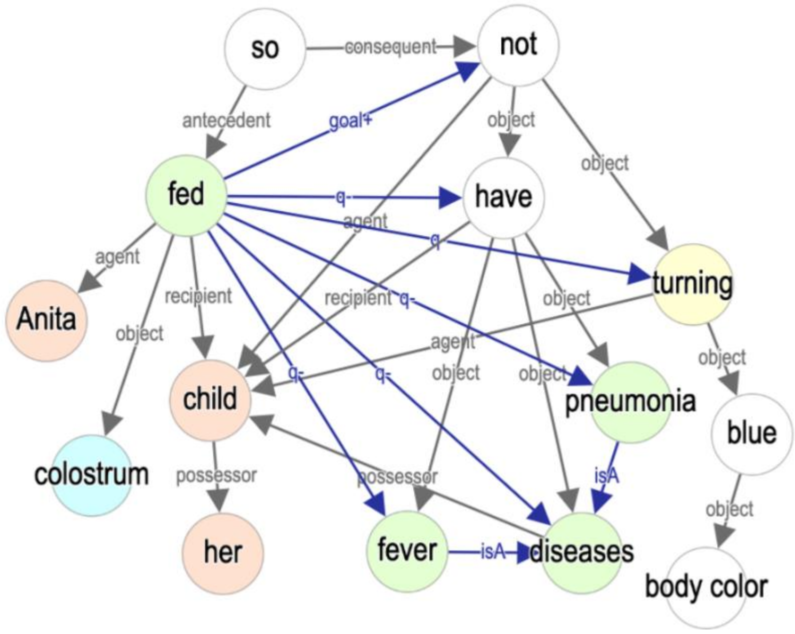

Figure 3. Semantic parse o “Anita fed colostrum to her child so that it does not have any diseases such as pneumonia, fever, turning body color blue.”

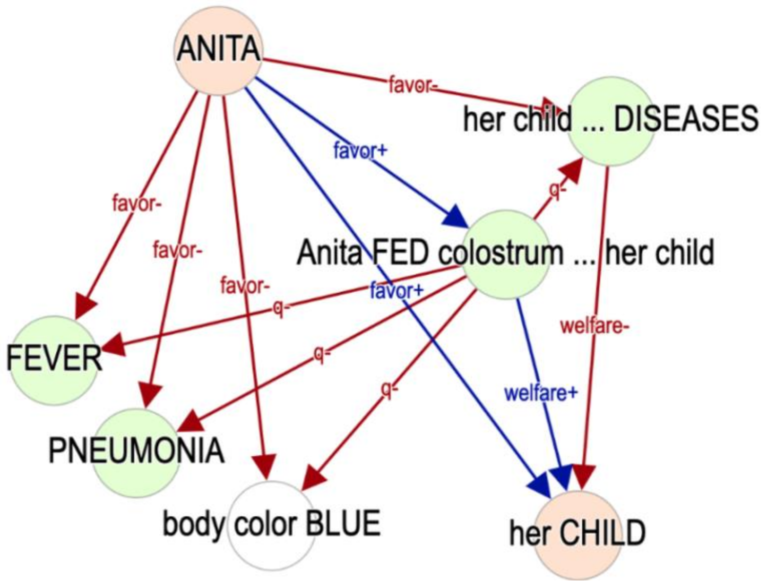

Figure 4. Causal model automatically refined from the semantic parse shown in Figure 2