Generative Tarot: Internet-Mining for the Collective Consciousness

Asiiah Song
University of California, Santa Cruz
1156 High Street
Santa Cruz, California 95064
julinas@ucsc.edu

Abstract

Tarot decks appeared as early as the 14th century. Once used as a game deck, today it is primarily used as a divination medium. One attempt to adapt the fortune-telling claims to modern technological pragmatism: the selection of card imagery are a representation of the *omni-homo*, or the every-human, or a collective consciousness, such that a card reading session becomes an interactive resurfacing of subconscious, but preexisting, knowledge. We suggest that perhaps when we create Tarot decks we are also creating a representation of the collective consciousness, in a manner of speaking. We provide a procedure of doing this creation from text-mining and take a look at the resulting artifact.

Introduction

Various accounts of the origin of Tarot assert that some form of the card deck appeared as early as the 14th century. While some pockets of Europe, such as Switzerland, still play Tarot as a game at least up to the late 20th century, most of the world encounters Tarot as a popular form of divination. (Dummet 1980) Many of us also encounter that divination in the context of play — we step into the magic circle and believe in the fortune-telling properties of the cards, an occasion that might occur more at festivals and parties than at the crossroads of life. There are also instances of using Tarot decks as a story-telling device. (Sullivan, Eladhari, and Cook 2018) An attempt at explaining the serious or semi-serious clientele for Tarot-based fortune-telling is that the selection of card imagery are a representation of the *omni-homo*, the every-human, or otherwise called a collective consciousness (this term used here as a metaphor for oft-occurring beliefs, concerns, and values of society), such that a card reading session becomes an interactive resurfacing of the client's subconscious but preexisting knowledge. (Aphek and Tobin 1989)

We present a project which generates novel Tarot decks by scraping the internet for archetypes and prevalent themes of our minds. The generation is modeled after classic Tarot decks such as Rider-Waite or Thoth (Michelsen 2005), which together form the base of the vast majority of existing decks. The goal is a process that creates a variation of decks when given different source material — different narrow sections of the internet — of a quality that is suitable for a divination session. The loftier motivation of the project is the question of how to operationalize humanness — how to store small representations of things a human might care about in knowledge graphs, which could potentially be used in AI agents for simulated artificial societies. We do not claim that these generated Tarot decks will be fair representations of the collective consciousness, but we suggest it could be an *operationalized* version, a watered-down model of the target.

Method

Structure of a Tarot Deck

We're trying to model after a typical Tarot deck. What does it look like? While there are variations to Tarot, in the following general description we write as if there is a single standard set and a standard method of interpretation. There are 78 cards. 22 of those are the named trump cards, the *major arcana*, such as *Wheel of Fortune*, *Fool*, or *Temperance*. The remaining 56, the *minor arcana*, are divided between four suits: *chalices*, *swords*, *wands*, and *pentacles*. Each set of suits cards contain the numbered cards 1 to 10 and four court cards: *page*, *knight*, *queen*, and *king*.

The semiotics of the language of a Tarot deck can be interpreted on several levels. Major arcana cards are interpreted primarily using imagery surrounding the subject matter of the card, identified by its name and also reflected in the illustration on the face of the card. Minor arcana cards are interpreted first by suit, then by numerology if it is a numbered card, and by socio- and biological traits of the court card if it is a court card, as well as imagery in the illustration. Interpretation of the illustrations are aided by colors and motifs. Aphek and Tobin (1989) posit that, the four suits are the most ordered and structured of those properties; colors, due to culturally and historically motivated interpretations, are semi-ordered; and the numerology-based interpretations are the least systematic.

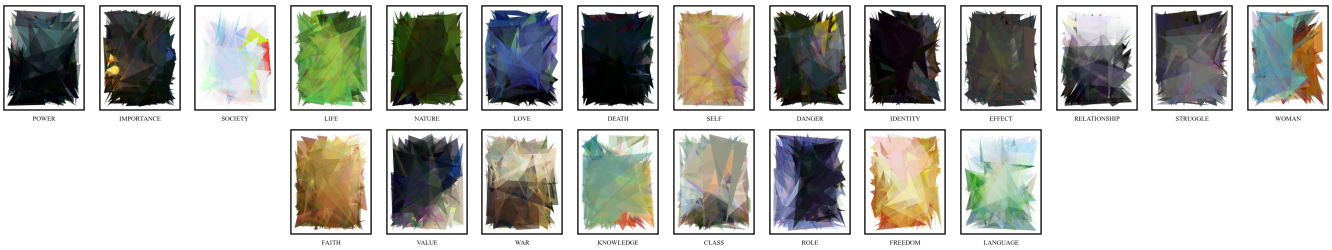

Figure 1: Generated major arcana. From left to right, row-wise: *0-Power, 1-Importance, 2-Society, 3-Life, 4-Nature, 5-Love, 6-Death, 7-Self, 8-Danger, 9-Identity, 10-Effect, 11-Relationship, 12-Struggle, 13-Woman, 14-Faith, 15-Value, 16-War, 17-Knowledge, 18-Class, 19-Role, 20-Freedom, 21-Language*.

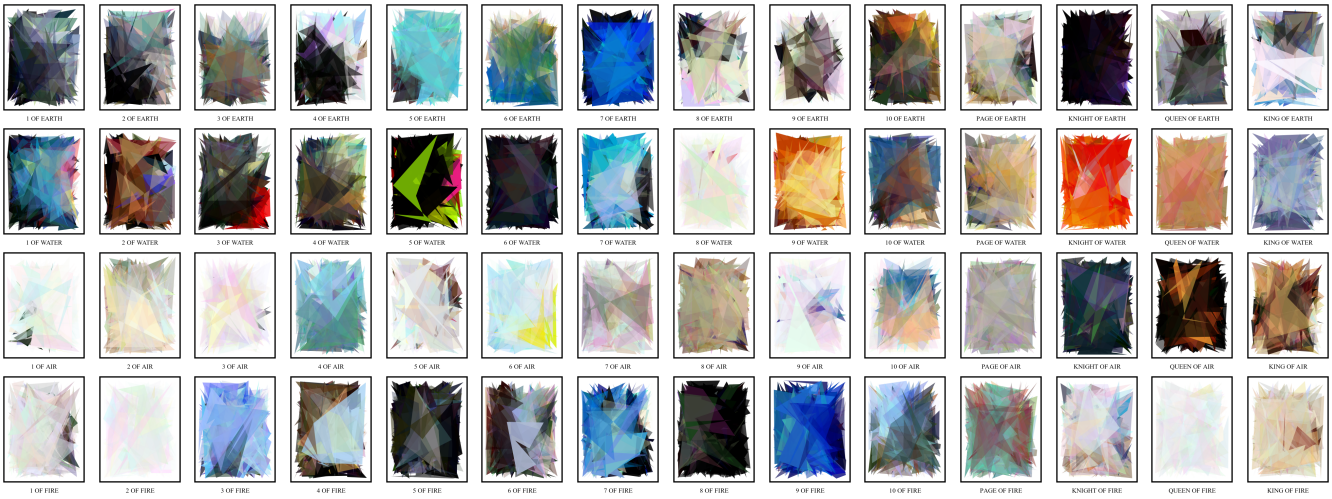

Figure 2: Generated minor arcana. Rows from top to bottom are: EARTH, WATER, AIR, and FIRE. Columns from left to right are: numbered cards 1 to 10, court cards page, knight, queen, and king.

Minor Arcana

1 of earth land, state, institution, reality, creation, manhood, initiation, country, humankind, world, humanity, earth, man, farming

2 of earth violence, force, intensity, strength, forcefulness

3 of earth way, dominance, agency, path, authority

4 of earth book, boy, son, word

5 of earth trouble, problem, difficulty

6 of earth face, side, expression

7 of earth escapism, flight, escape

Figure 3: A section of the generated booklet providing suggested readings.

Finding the Zeitgeist

The search for themes of our minds begins with a database that might contain a good range of that kind of information. In this project, we use the summarized theme headings of 588 literary works on SparkNotes.com as the source material. This is convenient in several ways: writing that get to be called literature generally deal with deep, reflective, human issues. This corpus is biased due to the choice of literary works included for review — the people who choose and the language, and the intended audience. Perhaps this bias is not a problem: it's hard to claim that human-made Tarot decks are not riddled with bias.

Scraped epithets are processed with SpaCy to extract thematic words from the text, and we pose them as ideas. Then, closely-related words are merged using WordNet's Leacock-Chodorow (LCh) similarity, which measures semantic similarity based on path distance between two *synsets* — a synset is essentially a definition of a word, which may have multiple definitions:

$$LCh = -\log \frac{shortest\_path(synset1, synset2)}{2 * depth}$$

A higher number indicates higher similarity. The shortest path here is simply how many connections removed one synset is from the other in WordNet. With this, we create a graph where each word is a node, and construct edges between nodes weighted with LCh similarity.

Mapping Ideas to Cards

We select the 22 most commonly-recurring words from our internet scrape and map them into trump cards. Then, for suit assignment, we apply sentiment analysis on the remaining ones on two axes (positivity/negativity and openness/closeness), designating a suit to each of the four quadrants. The suits correspond to the four European elements earth, water, air, and fire, following the example of existing Tarot decks. The correspondence of card suit to element is done using Leacock-Chodorow similarity between each element and each group of words.

To populate the 14 cards of each suit, we partition the graph into clusters with Networkx's built-in connected componenets subgraph function. Starting from the largest clusters (because we do not use all of them), we assign a random word in each cluster as the primary one of that cluster, and make it a card of the suit that the word belongs to. In an accompanying booklet to the deck, we record all words in that cluster as suggested interpretations of that card.

Each final card contains a framed illustration and a title under the frame. We obtain a source image for the illustration by retrieving a random photo from the Google Image Search API using the desired keywords. The image is arbitrarily cropped down to desired dimensions: the final goal image may be a single petal rather than a full flower, or the edge of a sign without including the main symbols. We run the genetic algorithm for approximating images with polygons on this source image. (Johansson 2008) 100 6-vertex polygons are used, with each mutation of a polygon changing the coordinates of the vertices or the RGBA value of the fill color. In each generation, the polygons are drawn in order onto the canvas, and the resulting RGB matrix of the image is used to calculate the fitness, assuming full opacity. The fitness function is the pixel-wise difference between the approximated image and the target image, returned as a percentage out of the maximum difference $width * height * 3\ channels * 255\ channel\ values$. We run the algorithm to a relatively low 85% fitness to obtain images that give a hint at the intended imagery without being explicit, thereby allowing freer interpretation — using abstraction to circumvent the difficulty of simulating human curation of dozens of imagery within a picture frame.

Results

Figure 1 and Figure 2 show the major and minor arcana cards of a complete generated deck. A different generation using the same source data would result in illustrations and in the text of the minor arcana. Half of the generated major arcana shown in Figure 1 has obvious correspondences, per primary readings, with the classical Rider-Waite deck (first in the pair is ours, second in the pair is Rider-Waite): 0-Power/8-Strength, 1-Importance/4-Emperor, 6-Death/8-Death, 7-Self/7-Hanged Man, 8-Danger/16-Tower, 9-Identity/9-Hermit, 11-Relationship/6-Lovers, 12-Struggle/7-Chariot, 13-Woman/3-Empress, 14-Faith/17-Star, and finally 17-Knowledge/01-Magician. Some of the major arcana illustrations have a dark tonality: that might say something about the color palette that is considered sombre or contemplative today. This is contrasted with the Rider-Waite deck, which considered yellow to be the color of calmness. Some illustrations are under-differentiated, which might be resolved by adding a curation step to the random source image retrieval from Google. The information density in the illustrations is also much lower than the classic Rider-Waite, which is a flaw of the simple generation procedure.

Figure 3 shows a section of the booklet generated for the same deck, providing suggested readings for the user. One thing missing from the generation procedure here is that the court cards are generated just like the numbered cards, without special care to include imagery of a person in either the illustration or the descriptions. However, knowing that they are court cards, a user can easily apply the basic reading of the cards to anthropomorphic connotations. While many Tarot decks have suggested readings for reversals (when a card is drawn vertically reversed), or advise that reversals should be interpreted as the opposite of the ordinary reading, we suggest, like the Thoth deck, not distinguishing reversals. This is due to the abstracted quality of the illustrations. It would certainly be an improvement if the automation can make the decision of whether or not to use reversals.

Figure 4 shows the fourth card in the earth suit in three decks: a traditional deck, a modern deck, and our generative deck.

The Python code is available at https://github.com/julinas/generative-tarot.

Figure 4: A comparison: 4 of DISKS in the Thoth deck, 4 of COINS in the Fountain deck, and 4 of EARTH in our deck.

Conclusion

While we cautiously claim that the resulting artifacts are usable in a Tarot reading session and have entertainment value, this project would be more complete with comparison and expressive range analysis using different source corpora, as well as user reviews from both serious users and recreational users. As to whether this was a successful operationalization of representation of human beliefs, concerns, and values, we think scraping pre-curated resources can be a way of extracting the raw material of beliefs, concerns, and values, but further weaving is needed to make something useful of it.

References

Aphek, E., and Tobin, Y. 1989. *The Semiotics of Fortune-Telling*. J. Benjamins Amsterdam & Philadelphia.

Dummet, M. 1980. *The Game of Tarot: From Ferrara to Salt Lake City*. Duckworth.

Johansson, R. 2008. Genetic programming: Evolution of mona lisa.

Michelsen, T. 2005. *The Complete Tarot Reader*. Llewellyn Worldwide.

Sullivan, A.; Eladhari, M. P.; and Cook, M. 2018. Tarot-based narrative generation. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*, 54. ACM.