Contents

Contents

Background Part I

1 The phenomenon of structural repetition

1.1 Repetition as an experience

We experience repetition in diverse domains: we discern shared narrative structures across novels, films, and music, and we uncover universal laws governing phenomena as distinct as ocean currents and atmospheric dynamics. Moreover, humans can abstract patterns of patterns, as seen in the application of category theory to mathematics and programming. These examples illustrate our remarkable ability to not only detect abstract repetitions but also to productively employ them in navigating both the physical and mental worlds.

Repeat, as we intuitively perceive it, is surprisingly nontrivial to formalize. As the term "repeat" is overloaded with multiple interpretations, we begin this section by clarify the *object* of repeat.

1.1.1 Repeated observations

Perhaps our simplest and most direct experience of repeat comes from the repetition of concrete observations. "Observation" refers to all kinds of direct sensory input. For instance, an ongoing fire siren corresponds to a repeat of a specific sound across time. The decorative motifs found in mosaic tiling are repeats of a visual symbol across space. The periodic motion of a pendulum contains repeats of a specific movement.

To put it formally:

$$[\![x_2 \; repeats \; x_1]\!]^1 = \{x_2 = x_1\} \qquad\qquad \text{(Approximation 1)}$$

Our first definition is simple and precise, and its mental processing of can be computationally modelled as pattern identification/matching.

However, exact pattern matching does not cover many of our usual experience of repeat. Consider a postman, when frustrated by his mundane work, complains about "doing the same thing" over and over again. Clearly, by "doing the same thing," our imaginary postman does

---

[1] We adopt the notation $[\![\ldots]\!]$ from formal semantics as a function that maps a expression to its truth condition

not mean it as repeating a list of concrete physical actions (e.g., turning right on the crossroad at 7:23 am, and stopping his car at the exact same spot). If the postman's job can be just explained as such, imagine the catastrophes caused by a mail robot programmed to do exactly that.

In order to make the postman's expression work, our model of repeat should allow some degrees of flexibility. For example, there are multiple ways to achieve the same task of "parking the car" depending on the situation at hand. What if the closest parking slot is occupied? What if there are kids walking by? What if there are construction works blocking the entire side of the street? "Just do the sensible things" would be a typical response from humans. In those scenarios, the observed actions would be adapted to include extra stepping on gas and break, checking the mirrors/traffic lights/pedestrians, looking for available parking slots. Then to our robot friend, in what ground can it understand these "sensible adaptions" of actions as being the "same"?

If we understand repeat based on the properties of the observations themselves, one way to extend our first approximation is to replace the equality in Approximation 1 by a similarity function. For example, one might say if two list of actions both ends with stopping the car and getting out the car, we define a high similarity among them.

1.1.2 Repeated abstractions

There are cases where our experience of repeat is not about the similarity of the observations at all. Imagine the cityscape of New York where we find repetitions of buildings. Each building can be very different from the image perspective. Base on the visual features alone, one might say the World Trade Center Transportation Hub (Figure 1.1) is closer to a dove than a building,[2] which is entirely valid as an *artistic* interpretation. However, few people would actually believe it is a real dove and would more likely to see it as a building in a shape of a dove.

The reason we perceive repeat in this circumstance is that we interpret these visuals as concrete instances of the same abstract concept—building—which we know its characterizing properties and functions. Thus, the core problem of the first definition is that it describes the phenomenon of the perceived repeat purely as a property of the observations themselves, without involving a mental understanding of the observations.

If we think of "mental understanding" as the ability to associate a specific token with its category/type, our next approximation of repeat would be "instantiations of the same category/type."

$$\llbracket x_2 \; \textit{repeats} \; x_1 \rrbracket = \{\mathrm{Type}(x_1) = \mathrm{Type}(x_2)\} \qquad \text{(Approximation 2)}$$

Categorization refers to the mental process of forming a system of tokens (observations) and their types/category (abstractions) and is considered a fundemetal human cognitive capac-

---

[2]The architect Santiago Calatrava use the appearance of a dove as a symbol for peace.

Figure 1.1: Flying Away I – Oculus by Calatrava New York. The photograph was shot in New York and was first published in August 2019.

ity. Computational mechanisms of human categorization has been theorized with varying representations of conceptual category. One idea is to use a "typical" observation as the representation of a conceptual category. One prominent cognitive theory that resembles such an idea is *Prototype theory* developed by psychologist Eleanor Rosch. Prototype theory describe the cognitive process of categorization as measuing the similarity between the observation with an ideal prototype of a concept in long term memory. In our postman example, it entails that we compare a specific realization of the action "parking the car" to a prototypical notion of parking the car in our memory (the prototype might differ from person to person). This can explain the postman's notion of "doing the same thing" by considering the specific implemetations of the actions as similar to an ideal version of it. It is worth noting that the ideal prototype may not actually exist in reality. For example a perfect circle does not exists in our physical world but it is still useful to categorize things base on ideal shapes. Continuing the same strand of thought, *Exemplar theory* propose that we do not just store one ideal prototype but all the instances of a concept, all of which may participate in the classification decision of an observation. These two notions of categorization manifests in data science as classification and clustering algorithms such as the Gaussian mixture model and the k-nearest neighborhood, which incorperates formal handling of the deviations from the representations (prototype or examplars) in memory.

Another approach to categorization is based on traits or affordance of entities (e.g. being able to run and carry objects) rather than their observable attributes (e.g. having four wheels). In the postman's example, we can categorize concrete actions by their intended effect on the world, for example, "inserting mail to mailbox" can be represented as any action that result in mail being in the mailbox. Such external characterization forms a useful abstraction that enables greater generality both in classification and in production. It also aligns closer with
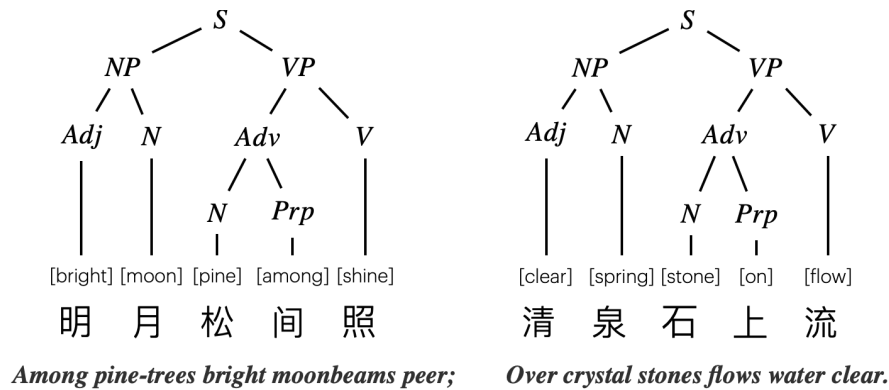
Figure 1.2: A poetry written by Wang Wei (699761 during Tang dynasty), translated by Xu Yuanchong, exhibiting parallel syntactic structure.

humans intuition in cases of innovation (e.g. imagining the next generation of transportation).

By shifting the focus from observations to their mental abstractions, we can explain the postman's notion of "doing the same thing" as multiple instantiations of abstract action "parking the car" or "inserting mail to mailbox" on different situation-encoding parameters such as "the location of the next closest parking slot" and "whether there are objects blocking the path". In the case of New York's cityscape, each appearance of the buildings can be thought of as an instantiated token from the type "building" using different configurations such as the building's structure, style, material, as well as spatial attributes and lighting environment such as the angle facing the viewer, light source and shadows, etc.

1.1.3   Repeated interpretations

Despite the generality of the approach, it may sometimes ignore too much information that are critical for our experience of repeat. Let's reflect on the common expression "the history repeats itself", not as a historical and philosophical statement, but as a result of our perception of repeat. Certainly, the notion of repeat is not about a concrete chain of events (based on our first approximation); and arguably, not about the overall abstraction of the historical events (based on our second approximation). It is rather about the pattern of how the smaller abstracted events (e.g. improved education, war, enactment of certain kind of law) contributes to the more abstracted events (advancement in technology and innovation, economic crisis, social inequality) which in turn leading to bigger events (e.g. global competitiveness,) and ultimately to the most abstracted description (e.g. the rise and fall of civilizations).

The same kind of inadequacy of our second approximation is revealed by an example of excerpt from a Chinese poem (Figure 1.2). One can notice that there are some correspondence between the structure of this two phrases. In terms of grouping, "[bright] [moon]" corresponds to "[clear] [spring]" and "[pine] [among] [shine]" corresponds to "[stone] [on] [flow]". Not only each of the corresponding character are from same syntactic category, the syntactic

Object of repeat	Corresponding cognitive ability
Concrete observations	Pattern matching
Abstractions	Categorization
Interpretations	Causal explanation

Table 1.1

structures of the two phrases are identical as demonstrated in Figure 1.2. Furthermore, the meaning of the corresponding words all aims to convey a feeling of serenity. In fact, these associations that one of the stylistic criteria for Tang dynasty is the usage parallel construction. Our percieved repeat (parallelism), even just on the syntactic level, can not be simply explained by that both phrases are noun phrases. If we substitute the second phrase with any other noun phrase, our previous experience of parallelism immediately disappear. If we explicitly factor in the interpretation about the observations, we are much closer to the kind of repeat shown in the Chinese poem example. Figure 1.2 reveals the syntactic structure (as one kind of interpretation) underlies the two parallel phrases. Compares to our second approximation (which would summarize both phrases as noun phrases), the parse tree provide the extra process connecting the observation to the abstraction. It is via such identical connections we experience a sense of repeat.

What is missing from our second approximation is the connection from the specific to the abstract. When describing repeat of this kind, we need to the able to refer to an underlying interpretation, a group of hierarchically organize relations linking layers abstractions to the concrete tokens.

$$[\![x_2 \ \textit{repeats} \ x_1]\!] = \{\exists R \ X_1 \ X_2, \quad X_1 R x_1 \wedge X_2 R x_2\} \qquad \text{(Approximation 3)}$$

1.2   Structural Repeat

Our experience of repeat can be enriched and deepened by structure. The previous three approximations for repetition captures increasingly nuanced experience of repeat via domain specific structure (see Table 1.1). In the first approximation, there is no structured knowledge or theory about the specific domain other than a notion of equality or similarity. In the second approximation, we assume an ontology of the domain that organize observations by their abstractions via a system of tokens and types. In the third approximation, we assume a casual theory which adds on top of the ontology, a set of primitive relations linking entities and their abstractions.[3]

In an informal way, "structure" describes "how things can be put together" or "how things can be decomposed" in a domain. When describing "how things can be put together", a good

---

[3]In particular, formal grammars can be seen as an instance of a causal theory where the ontology is the system of non-terminal and terminal symbols, and the primitive relations are the production rules.

system would have a few rules but can explain potentially infinite combinations. To achieve this, one need to be able to describe combinations on the level of abstractions/types instead of the specific instances as it gives rise to the ability to generalize. In other words, one need an ontology of the domain.

In natural language, words are abstracted into their *syntactic categories*, whose valid combination is described by the *syntax*. These two components specifies a causal theory of grammatically correct sentences.

In action planning, actions are abstracted into their *intended goals*, whose valid combination is described by the *inferences rules in logic*. These two components specifies a causal theory of goal-oriented actions.

In music, concrete notes are abstracted into *harmonies and voices or tones*,[4] whose valid combination is described by rules of *tonal structure*. These two component specifies a causal theory of tonal polyphony.

The main theme of this thesis is to understand repetition respecting the underlying causal structure. To be able to computationally model causal theory in a way that is domain general, we need to adopt a uniform representation of a causal theory. Let us examine three examples.

1.2.1   A causal structure of Chinese poetry

In this example, the real observable entities are the words "[bright]","[moon]", etc. The job of the causal theory is to explain the occurrence of these specific sequence words. The Chinese grammar fulfills the role of a causal theory in the following way. First an ontology of the entities (words) are specified in two parts: (1) the space of types, represented by syntactic categories such as nouns, verb, etc, and (2) the association from tokens to types[5]. Second a finite list of relations associating a type with n tokens or types, encoding how each syntactic category could be elaborated into multiple categories[6]. For example, the relation $R1$ says that a sentence and be constructed as a noun phrase followed by a verb phrase, which asserts the grammatical validity of combining "[bright] [moon]" with "[pine] [among] [shine]"[7].

With a causal theory in place, a causal explanation is made concrete as a composition of the domain specific relations (Figure. 1.3), representing an interpretation of the observations. At this point one might wonder why we do not use a tree use node are types (as the ones in Figure 1.2). The answer is that when the relations have multiple instantiations (if it involves variables for instance), a tree of types could not capture the sameness of the underlying relation. However, a tree whose node are relation explicitly represents the structure of the

---

[4]Notes without temporal attributes

[5]This corresponds to non-terminal symbols and termination rules in formal grammar

[6]These relations corresponds to a subset of the production rules in formall grammar where the right-hand side does not contain non-terminals.

[7]The second part is valid according to the Chinese grammar and is invalid with the English grammar.

8

Ontology	$\text{types} = \left\{ \begin{array}{c} S \\ NP \\ N \\ Adj \\ Adv \\ Prp \end{array} \right\}$	$\text{tokens} = \left\{ \begin{array}{r} \text{bright, clear} \\ \text{moon, pine, spring, stone} \\ \text{among, on} \\ \text{shine, flow} \end{array} \right\}$
Relations	$\dfrac{S}{NP \quad VP}\text{R}_1 \qquad \dfrac{NP}{Adj \quad N}\text{R}_2 \qquad \dfrac{VP}{Adv \quad V}\text{R}_3$ $\dfrac{Adv}{N \quad Prp}\text{R}_4 \qquad \dfrac{N}{\text{moom} \mid \text{pine} \mid \text{spring} \mid \text{stone}}\text{TERM}_N$ $\dfrac{Adj}{\text{bright} \mid \text{clear}}\text{TERM}_{Adj} \qquad \dfrac{Prp}{\text{among} \mid \text{on}}\text{TERM}_{Prp} \qquad \dfrac{V}{\text{shine} \mid \text{flow}}\text{TERM}_V$	

Table 1.2: A simplified casual theory in the domain of language (the Chinese poem example)

interpretation. If we understand the composition of relations as a generative program, then the type tree corresponds to the input and output of our program after each function call, whereas the relation tree corresponds to the program itself. Our next example in action planning demonstrates the additional expressive power of relation tree when a relation have multiple instantiations.

1.2.2 A causal structure of action planning for making coffee

Imagine John performed the following list of actions: he measures 18 grams of coffee ground, put these measured coffee ground into a coffee maker, he then measures 150 ml of water, and put the measured water into the coffee maker, then he activates the coffee machine. We are first interested in how we can understand John's action as a coherent realization of a goal, which is answered by a causal theory of the domain. We further ask how can be capture our percieved repeat within this list of actions, which is answered by infering the shared component within a causal explaination of the actions.

Similar to the Chinese poem example, the ontology of action can be defined in two parts (1) a space of types represent goals and (2) a mapping from tokens to types representing how an action can be understood as to achieve a goal. To make things simple, goals are interpreted as changes in the state of the world, denoted by an arrow from state before the action to the state after the action. For example, the overarching goal in this example is that John want to transform his state from "having coffee bean and water supply" into "having a cup of coffee".
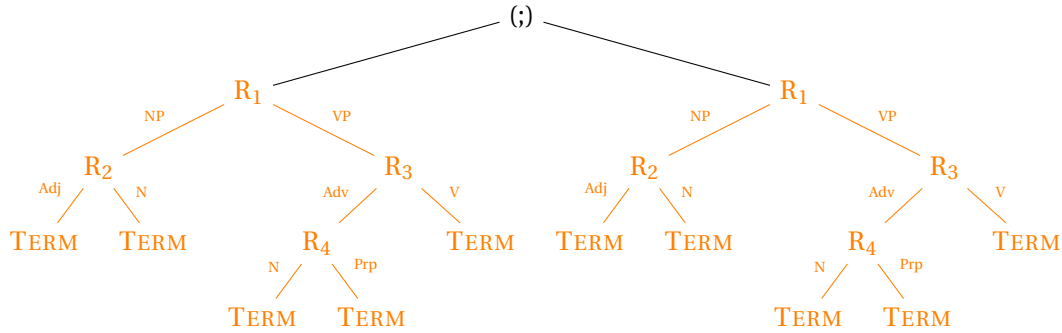
Figure 1.3: The composition of domain specific relations as repeated object in the Chinese poetry example

Goals can be combined in two general and natural ways, the first is sequential composition (chaining): if we know how to accomplish $x \to z$ and $z \to y$ we can accomplish the goal $x \to y$ by concatenating the two actions (also feeding the result of the action to the next action). The second is parallel composition, if we can accomplish $x_1 \to y_1$ and $x_2 \to y_2$, we can also accomplish $(x_1, x_2) \to (y_1, y_2)$. The tuple notation denotes the conjunction of states.

These rules about how goal can be decomposed forms the space of relations in our action planning domain.

1.3   Cognitive principles underlying syntactic repetition

### 1.3.1   Compossibility

1.3.2   Efficient reuse of computations

1.3.3   Minimizing description length

Ontology

$$\text{types} = \{s_1 \rightarrow s_2 \mid s_1, s_2 \in \text{WorldState}\}$$
$$\text{WorldState} = \{\text{Having } q \mid q \in \text{Quantified}\}$$
$$\cup \{x \text{ in machine} \mid x \in \text{Quantified}\}$$
$$\cup \text{WorldState} \times \text{WorldState}$$
$$\text{Quantified} = \{x \text{ amount of } y \mid x \in \mathbb{R} \times \text{Unit}, y \in \text{Entity}\}$$
$$\text{Unit} = \{\text{gram, milliliter, cup}\}$$
$$\text{Entitiy} = \{\text{coffee ground, coffee, water}\}$$

$$\text{tokens} = \left\{ \begin{array}{c} \textbf{put } x \textbf{ in machine} \\ \textbf{measure } x \textbf{ amount of } y \\ \textbf{activate the machine} \end{array} \right\}$$

Relations

$$\frac{x \rightarrow y}{x \rightarrow z \quad z \rightarrow y}\text{CHAIN} \qquad \frac{(x_1, x_2) \rightarrow (y_1, y_2)}{x_1 \rightarrow y_1 \quad x_2 \rightarrow y_2}\text{AND}$$

$$\frac{(x \rightarrow x \text{ in machine})}{\textbf{put } x \textbf{ in machine}}\text{PUT} \qquad \frac{\text{having supply of } x \rightarrow \text{having } y \text{ amount of } x}{\textbf{measure } x \textbf{ amount of } y}\text{MEASURE}$$

$$\frac{\left(\begin{array}{c} 18 \text{ g coffee ground in machine} \\ 150 \text{ ml water in machine} \end{array} \right) \rightarrow \text{Having a cup of coffee}}{\textbf{activate the machine}}\text{MACHINE}$$

Table 1.3: A simplified casual theory in the domain of goal-oriented actions (the making coffee example)
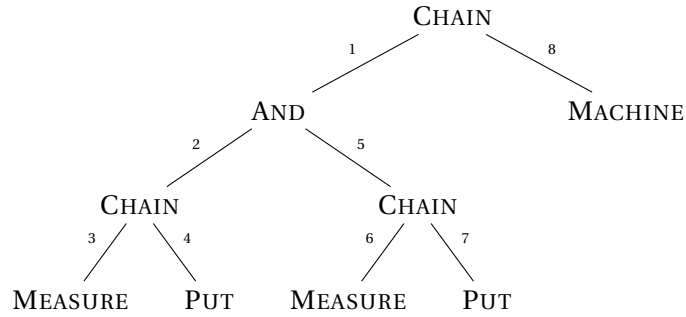
$$1 = \begin{pmatrix} \text{Having supply of coffee ground} \\ \text{Having supply of water} \end{pmatrix} \rightarrow \begin{pmatrix} \text{18 g coffee ground in machine} \\ \text{150 ml water in machine} \end{pmatrix}$$

$2 = \text{Having supply of coffee ground} \rightarrow \text{18 g coffee ground in machine}$

$3 = \text{Having supply of coffee ground} \rightarrow \text{Having 18 g coffee ground}$

$4 = \text{Having 18 g coffee ground} \rightarrow \text{18 g coffee ground in machine}$

$5 = \text{Having supply of water} \rightarrow \text{150 ml water in machine}$

$6 = \text{Having supply of water} \rightarrow \text{Having 150 ml water}$

$7 = \text{Having 150 ml water} \rightarrow \text{150 ml water in machine}$

$$8 = \begin{pmatrix} \text{18 g coffee ground in machine} \\ \text{150 ml water in machine} \end{pmatrix} \rightarrow \text{Having 1 cup of coffee}$$

$$\text{goal} = \begin{pmatrix} \text{Having supply of coffee ground} \\ \text{Having supply of water} \end{pmatrix} \rightarrow \text{Having 1 cup of coffee}$$

Figure 1.4: The composition of domain specific relations as repeated object in the making coffee example

2 Syntactic repetition in tonal music

2.1 Hierarchical structure in Music

2.1.1 Form, rhythm, harmony

2.2 Repetition in Music

2.2.1 Exact vs varied

2.2.2 Syntactic vs non-syntactic

2.2.3 Differentiating reuse and repeat

3 Bayesian model-based reasoning

3.1 Generative models

3.2 Functional Probabilistic Programming

3.2.1 Markov category and probability monads

3.2.2 Synthetic probability theory and Quasi-Borel space

3.2.3 Bayesian non-parametrics

Template Grammar Part II

Computational Experiments Part III

4 Pattern discovery in Jazz Harmony Tree Bank

4.1 Quantitative results and discussion

4.2 Qualitative analysis

5 The plausibility of minimal-sized Template as Jazz harmonic analysis

6 Inducing Template grammar from Jazz chord progressions

6.1   Method 1: Non-parametric Bayesian inference

6.1.1   Pitman-Yor process prior

6.2   Method 2: A heuristic based on tree compression

6.2.1   Straight-line tree grammar

6.3   Comparison between the two methods

7 Contributions and conclusions