RecipeScape: Mining and Analyzing Diverse Processes in Cooking Recipes

Minsuk Chang
School of Computing
KAIST
minsuk@kaist.ac.kr

Juho Kim
School of Computing
KAIST
juhokim@kaist.ac.kr

Vivian M. Hare
Computer Science Department
Stanford University
vhare@stanford.edu

Maneesh Agrawala
Computer Science Department
Stanford University
maneesh@cs.stanford.edu

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053118

Abstract
In culture analytics, it is important to ask fundamental questions that address salient characteristics of collective human behavior [11, 1, 17] . This paper explores how analyzing cooking recipes in aggregate and at scale identifies these characteristics in the cooking culture, and answer fundamental questions like ”what makes a chocolate chip cookie a chocolate chip cookie?”. Aspiring cooks, professional chefs and cooking hobbyists share their recipes online resulting in thousands of different procedural instructions towards a shared goal. However, existing approaches focus merely on analysis at the ingredient level, for example, extracting ingredient information from individual recipes. We introduce RecipeScape, a prototype interface which supports visually querying, browsing and comparing cooking recipes at scale. We also present the underlying computational pipeline of RecipeScape that scrapes recipes online, extracts their ingredient and instruction information, constructs a graphical representation, and computes similarity between pairs of recipes.

Author Keywords
Culture Analytics; Interactive Data Mining; Mining at Scale; Recipe Mining; Naturally Crowdsourced Data

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

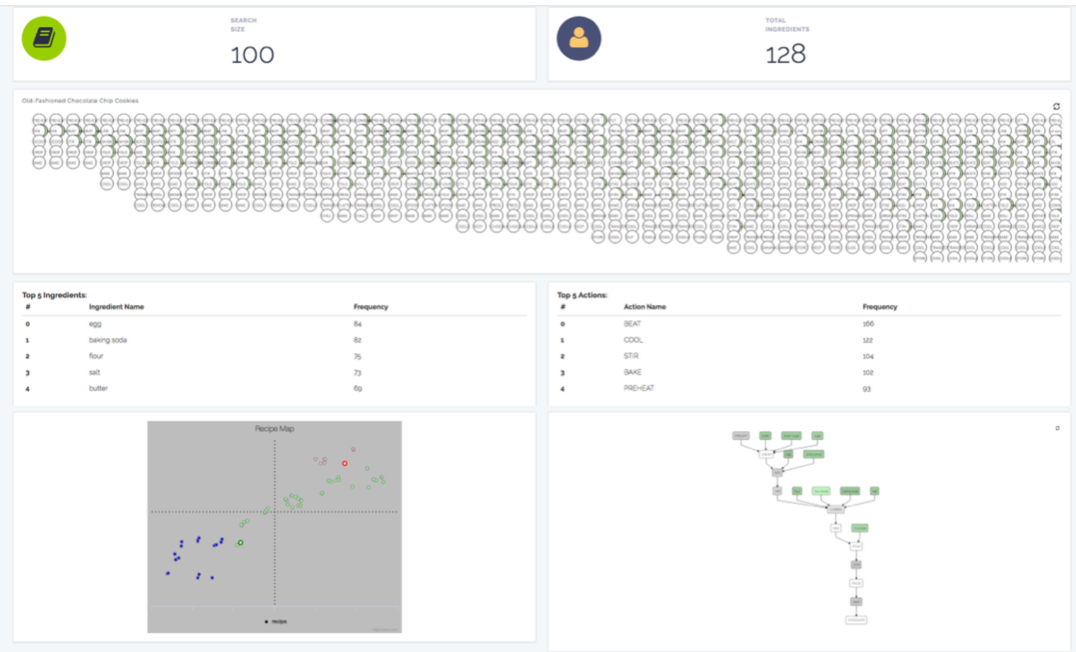

Figure 1: RecipeScape Dashboard Interface

Introduction

"The proportion of ingredients is important, but the final result is also a matter of how you put them together." [16] said Alain Ducasse, a world class chef who is currently decorated with 18 Michelin stars. Personal preferences, tips and know-hows in cooking are well represented in the thousands of recipes that aspiring cooks, professional chefs and cooking hobbyists share online. As a result, thousands of recipes are available across the internet even for a single a dish like chocolate chip cookie, roast turkey and meatball pasta.

In essence, these recipes are naturally crowdsourced instructions for a shared goal, like making a chocolate chip cookie. Comparing, combining and investigating these recipes as processes at scale could unveil the collective human knowledge and human practices around cooking, and increase our understanding of behaviors and dimensions in culture analytics [14].

While many websites offer detailed information about individual recipes, they do not offer comprehensive analytics of a specific dish. Most services are limited to searching specific ingredients or comparing differences in ingredients, providing categorization based on the type of cuisine or originating culture, and simple recommendations based on *like*s of other users. Existing research focuses on extracting ingredient information [4], or modelling a single recipe into a sequential structure [7, 13]. But to the best of our knowledge, little work has compared multiple recipes as procedural workflows, or identified common steps and characteristics across them in aggregate.

This paper investigates to uncover underlying insights into cooking by examining these recipes in aggregate and at scale. If we can visualize hundreds of chocolate chip cookie recipes in aggregate, and if we can look at the embedded processes of the recipe instructions at once, can we discover common trends and distributions of chocolate chip cookie recipes? If we can compare and analyze these instructions in more dimensions than just ingredients, can we find out what the median recipe looks like, and what components make the median recipe? Can we find out what most creative variants are? Can we find out the simplest and the most complex recipes? Can we infer what makes a choco-

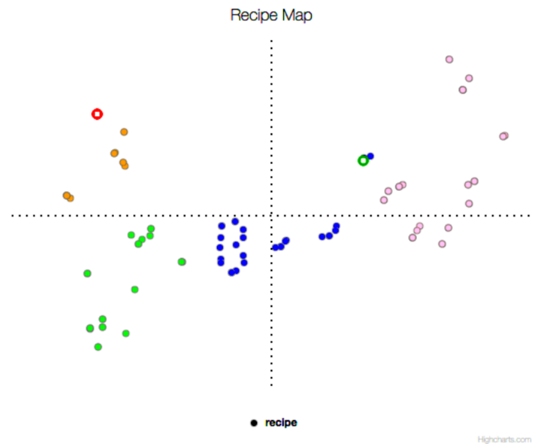

Figure 2: RecipeMap

late chip cookie a chocolate chip cookie? Ultimately, can we find out fundamentally different ways of making chocolate chip cookies?

In order to answer these questions, we present RecipeScape (Figure 1), a prototype interface that enables interactive browsing, comparison, and visualization of recipes at scale. RecipeScape is powered by a novel computational pipeline that scrapes recipes from online websites, extracts their ingredients and instructions information, translates them into a graphical representation, and computes similarities between all pairs of recipes.

RecipeScape

In RecipeScape, we provide two main interactive visualizations, RecipeMap and RecipeDeck, which support aggregate macro views of the recipes. From the two visualizations, users can perform a more fine-grained level of analysis using a pairwise comparison view, namely RecipeDuel, and by examining individual recipes.

RecipeMap

In order to support queries for the median recipe or outlier recipes, and to provide the overall distribution of recipes for a specific dish, we provide RecipeMap (Figure 2). Each point on the map is a recipe, and their distance from one another represents their structural similarity. The median recipe can be recovered from selecting a recipe in the middle of the RecipeMap. The outliers are located in the sides and corners. Recipes are clustered based on the structural similarity, and different clusters are presented in different colors. The details of the similarity metric will be discussed in the pipeline section. This two dimensional plot is inspired by scatterplot to provide a familiar and commonly used visualization for presenting a distribution.

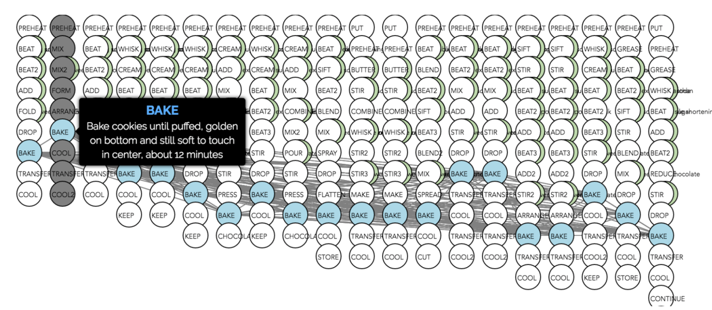

Figure 3: RecipeDeck

RecipeDeck

RecipeDeck (Figure 3) provides a layered comparison view of recipe processes which allows the user to understand how a specific cooking action is used in comparison to other recipes in the same cluster. Individual recipes are represented as a list of cooking actions and the lists are horizontally stacked. The cooking actions that involve ingredients have green shades. As shown in Figure 3, if the user selects the "bake" node in the second recipe, all "bake" nodes in the other recipes are highlighted. This interaction allows the users to conveniently observe where a certain cooking action like "bake" happens in all other recipes, and it enables the discovery of trends and patterns of diverse cooking actions.

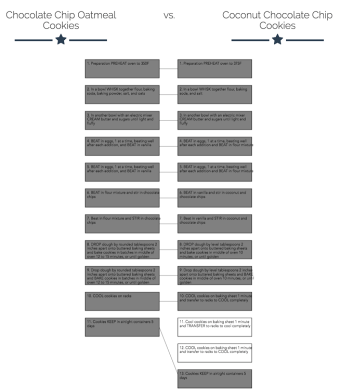

Figure 4: Recipes with High Similarity

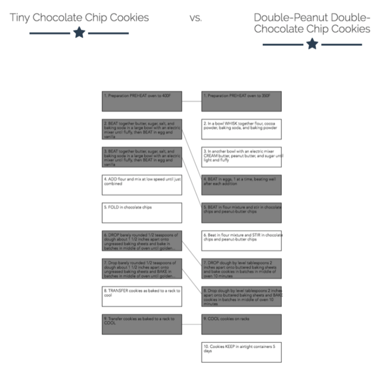

Figure 5: Recipes with Medium Similarity

RecipeDuel

From the RecipeMap, the user can select any two recipes for a pairwise comparison view, namely RecipeDuel. This side-by-side comparison lets users visually understand the similarities and differences in the process. The two recipes are presented in a union graph [8], where steps that are similar in both recipes are highlighted in gray and connected by an edge. Example views of two similar recipes and two distant recipes are shown in Figure 4 and Figure 5, respctively. The union graph makes it easy to grasp the degree of pairwise similarities and at the same time offers a guide for where to examine for procedural differences.

Pipeline

Representing recipes as trees is a logical decision because of the sequential nature of recipes, i.e., multiple ingredients are connected to a cooking action and cooking actions are also connected to each other. It has a critical advantage because user level analytics tasks can be achieved by utilizing various graph algorithms. For example, we can calculate similarities between two recipes using graph edit distances. In this section, we present the underlying pipeline (Figure 6) of RecipeScape for constructing graphical representations of recipes and obtaining similarity metrics by highlighting the data gathering, parsing and similarity comparison steps.

Data Gathering

In the data gathering step, first we query a dish and crawl all recipe results from recipe websites that use the schema.org's Recipe item type [1]. Then we parse the collected recipe documents for ingredients and instructions and extract the corresponding information into raw texts.

Parsing

Researchers have been exploring natural language processing and machine learning approaches to parsing recipes and to modeling the sequential information. Two of the most relevant works are SIMMR [4] and Mise en Place [7]. SIMMR is an NLP parser which converts recipes in the CURD [18] format into a tree representation. SIMMR requires recipes to be in the custom format, hence falls short in dealing with naturally crowdsourced recipes with diverse expressions and sentence structures. Mise en Place uses a Hidden Markov Model to model instructions of recipes

[1] http://schema.org/Recipe

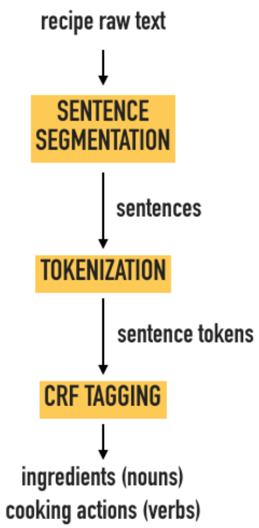

Figure 6: Pipeline

and uses an Expectation-Maximization (EM) algorithm to recover the structure of sequences of actions from natural language recipe texts. However, the probabilistic graphical models [7] describe individual recipes with emphasis on communicating the uncertainty of the parsing prediction, making them not suitable as the underlying representation for making pairwise comparisons, browsing, or trend analysis.

For the parsing step, we first apply a pre-trained conditional random field (CRF) [2] to tag relevant tokens with cooking actions and ingredients. To obtain the training set for the CRF, we designed a web-based annotation interface for crowdsourcing tags (Figure 7). The user is given a sentence in a recipe and tags tokens as cooking actions, cooking ingredients, or cooking tools. We use the token - tag pairs as labelled training examples.

Our pipeline then generates a tree representation for each recipe by lining cooking verbs in the order they appear in the recipe text. After the sequence of cooking verbs are lined, the associated ingredients to each cooking verb are attached as child nodes (Figure 8). We found this method

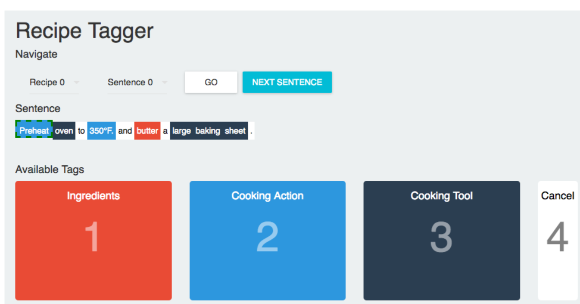

Figure 7: Recipe Tagger Interface

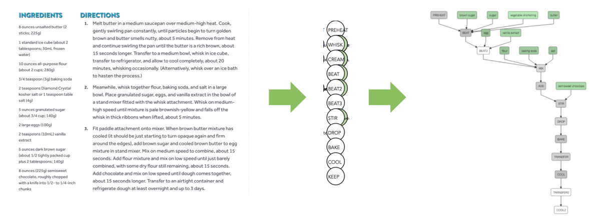

Figure 8: Parsing Process

simple but effective because it takes advantage of the sequential nature of the recipe instructions.

Similarity Comparison
In order to obtain the similarities between the recipes, we topologically sort the generated trees, and compare the median topologies of each recipe with one another using the weighted Levenshtein distance [10]. This similarity information is stored in a matrix.

The similarity matrix is then converted into a plottable vector using T-SNE [12], which is a dimensionality reduction technique particularly suitable for visualizing high dimensional data due to its capability of preserving the relationships between data points in the embedding space. And these recipes are plotted on RecipeMap. In order to highlight structurally different clusters of recipes, we use the K-Means clustering algorithm [3].

Case Study

For a proof of concept, we used RecipeScape to examine 487 chocolate chip cookie recipes. We were able to make interesting observations. From the RecipeMap, we were able to uncover the median recipe from the cluster with most recipes; ingredients related to making cream are added first, then flour is added and mixed to the cream to make the dough, then chocolate chips are added last, and finally the cookies are baked. Also recipes that are very different from the median recipes are found (Figure 9. We also noticed one cluster of recipes always involves oatmeal as ingredients.

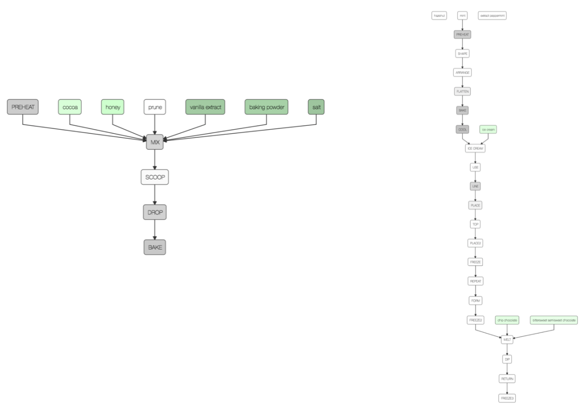

Figure 9: Two recipes away from the median.

From RecipeDeck, we were able to identify cooking actions that have high dependency by observing pairs of actions that always appear together such as ”preheat” and ”bake” across all recipes. We also discovered the exchange of orderings between the steps ”add”and ”stir”, and by referring to RecipeDuel, we noticed ”add” followed by ”stir” yields different texture in the cookie from ”add” and ”mix”.

Conclusion and Future Work

In this paper, we introduced RecipeScape, our prototype tool for recipe comparison and analytics in aggregate and at scale. We also introduced a computational pipeline that retrieves the procedural structure of recipes and provides similarity information between recipes. Through RecipeScape, we have shown that analyzing cooking recipes in aggregate and at scale is a promising channel that provides initial understanding of collective knowledge, and practices around cooking.

There are multiple directions for future work. We plan to support additional queries, such as determining ordering constraints in a sequence of actions, and examining the duration of processes. We also plan to further improve similarity metric to be able to catch the subtle but interesting semantic differences, and explore more robust representations and similarity comparison methods. Some candidates are measuring text similarities based on vector space embeddings [9, 15], characteristic vectors [5] and tree edit distance [6]. Currently RecipeScape emphasizes the process-level analysis, but we also plan to explore the relationships between the process dynamics and the outcomes.

RecipeScape is an analytics tool in its current form, but we believe our pipeline can further support learning and real-time planning for cooks. We plan to augment individual recipes with collective knowledge and generate an interactive visualization that the user can navigate through while cooking. We would also like to extend to examining video based recipes and how-to cooking videos in aggregate to discover more context rich trends and patterns into the embedded processes in cooking.

Acknowledgements

This work was supported by ICT R&D program of MSIP/IITP [R7124-16-0004, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding]. We thank Brown Institute for Media Innovation at Stanford University for their valuable comments and support in this work.

References

[1] Peter Sheridan Dodds, Kameron Decker Harris, Isabel M Kloumann, Catherine A Bliss, and Christopher M Danforth. 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PloS one* 6, 12 (2011), e26752.

[2] Erica Greene. 2015. Extracting Structured Data From Recipes Using Conditional Random Fields. https://open.blogs.nytimes.com/2015/04/09/extracting-structured-data-from-recipes-using-conditional-random-fields/. (2015). [Online; Published 9-Apr-2015].

[3] John A Hartigan and Manchek A Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 1 (1979), 100–108.

[4] Jermsak Jermsurawong and Nizar Habash. 2015. Predicting the Structure of Cooking Recipes. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 781–786.

[5] Lingxiao Jiang, Ghassan Misherghi, Zhendong Su, and Stephane Glondu. 2007. Deckard: Scalable and accurate tree-based detection of code clones. In *Proceedings of the 29th international conference on Software Engineering*. IEEE Computer Society, 96–105.

[6] Tao Jiang, Lusheng Wang, and Kaizhong Zhang. 1995. Alignment of treesâĂŤan alternative to tree edit. *Theoretical Computer Science* 143, 1 (1995), 137–148.

[7] Chloé Kiddon, Ganesa Thandavam Ponnuraj, Luke Zettlemoyer, and Yejin Choi. 2015. Mise en Place: Unsupervised Interpretation of Instructional Recipes. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 982–992.

[8] Nicholas Kong, Tovi Grossman, Björn Hartmann, Maneesh Agrawala, and George Fitzmaurice. 2012. Delta: a tool for representing and comparing workflows. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1027–1036.

[9] Quoc V Le and Tomas Mikolov. 2014. Distributed Representations of Sentences and Documents.. In *ICML*, Vol. 14. 1188–1196.

[10] Vladimir I Levenshtein. 1966. Binary codes capable of correcting deletions, insertions and reversals. In *Soviet physics doklady*, Vol. 10. 707.

[11] Stanley Lieberson and Joel Horwich. 2008. Implication analysis: a pragmatic proposal for linking theory and data in the social sciences. *Sociological Methodology* 38, 1 (2008), 1–50.

[12] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9, Nov (2008), 2579–2605.

[13] Jonathan Malmaud, Jonathan Huang, Vivek Rathod, Nick Johnston, Andrew Rabinovich, and Kevin Murphy. 2015. What's Cookin'? Interpreting Cooking Videos using Text, Speech and Vision. *arXiv preprint arXiv:1503.01558* (2015).

[14] Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K Gray, Joseph P Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, and others. 2011. Quantitative analysis of culture using millions of digitized books. *science* 331, 6014 (2011), 176–182.

[15] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. 3111–3119.

[16] Samuel Muston. 2013. Alain and the chocolate factory: How a celebrated French chef became an unlikely Willy Wonka. http://www.independent.co.uk/life-style/food-and-drink/features/alain-and-the-chocolate-factory-how-a-celebrated-french-chef-became-an-unlikely-willy-wonka-8790417.html. (2013). [Online; Published 30-Aug-2013].

[17] Melanie Swan. 2013. The quantified self: Fundamental disruption in big data science and biological discovery. *Big Data* 1, 2 (2013), 85–99.

[18] Dan Tasse and Noah A Smith. 2008. *SOUR CREAM: Toward semantic processing of recipes*. Technical Report. Technical Report CMU-LTI-08-005, Carnegie Mellon University, Pittsburgh, PA.