RC25470 (WAT1405-061) May 28, 2014
Computer Science

IBM Research Report

New Developments in Culinary Computational Creativity

Nan Shao, Pavankumar Murali, Anshul Sheopuri
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
USA

New Developments in Culinary Computational Creativity

Nan Shao, Pavankumar Murali, Anshul Sheopuri
IBM TJ Watson Research Center, Yorktown Heights, NY

Abstract

In this paper, we report developments in the evaluation and generation processes in culinary computational creativity. In particular, we explore the personalization aspect of the quality and novelty assessment of newly created recipes. In addition, we argue that evaluation should be a part of the generation process and propose an optimization-based approach for the recipe creation problem. The experimental results show a more than 41% lift in the objective evaluation metrics when compared to a sampling approach to recipe creation.

1 Introduction

"My children have a preference for meat. How do I create a healthy dish that will be enjoyed by them?" Can a computer help parents with such questions?

Culinary domain is a new area for computational creativity, although "made up a recipe" has been listed as one of the 100 creative activities on human creativity rating questionnaire developed by Torrance more than 50 years ago (Sawyer 2012). (Morris et al. 2012) discussed recipe creation restricted to soups, stews and chili. (Varshney et al. 2013) discussed evaluation (work product assessor) motivated by neural, sensory and psychological aspects of human flavor perception, and proposed models for culinary computational creativity system.

To answer questions like the one listed above, we consider two aspects of the problem: the personalization of dish evaluation and the optimization of dish quality and novelty in a combinatorially complex creativity space. Our contributions to the culinary domain are as follows. First, creativity is only meaningful in the presence of a human audience or evaluator (Wiggins 2006), and humans are inherently different; therefore we explore the personalization aspect of the evaluation metric for a creative artifact. In particular, we consider flavor preference and novelty evaluation of a newly created recipe. Second, we consider evaluation as part of the generation/search process and provide an optimization-based approach for the recipe creation problem. For the latter, we draw inspiration from the search mechanism that (Wiggins 2006) proposed on moving through the complex conceptual space. We hypothesize that our proposed methodological framework can be extended to other creative endeavors as well.

2 Personalization in Culinary Creation

We now turn to detailing a tractable approach for assessing personalized flavor preference and novelty. The approach is motivated by the human flavor perception science, technology to draw information from the web, and the work in (Varshney et al. 2013).

2.1 Flavor Preference

Flavor enhancement, balance and substitution are choices that we make to live a healthy life. Often, we may want to enhance the flavor of our favorite ingredient. However, we may need to balance the flavor of healthy but not tasteful ingredients. Moreover, we may want to substitute red meat with a plant-based product to meet the dietary constraint and, at the same time, not lose the meaty flavor. In our work, we propose a methodology to address these personalized flavor preferences in a computational creativity system.

Knowledge of how humans perceive flavors is necessary to build a system that accurately estimates a human's evaluation for creativity. For this reason, (Varshney et al. 2013) proposed a model for pleasantness which correlates olfactory pleasantness with its constituent ingredients and flavor compounds in those ingredients based on recent olfactory pleasantness study (Haddad et al. 2010; Khan et al. 2007). The smell of food is a key contributor to flavor perception, which is, in turn, a property of the chemical compounds contained in the ingredients (Burdock 2009; Shepherd 2006). Therefore a tractable step towards a data-driven model for flavor enhancement, balance and substitution is a model for odor similarity. For example, we could enhance the flavor of a featured ingredient by adding other foods with perceptually similar odors.

Recent work has shown that perceptual similarity of odorant-mixtures can be predicted (Snitz et al. 2013). Consistent with the synthetic brain processing mechanism in olfaction, human perception groups many mono-molecular components into singular unified percept. Each odorant-mixture is modeled as a single vector made up of the structural and physicochemical descriptors of the mixture. The angle distance between two vectors is a meaningful predictor of the perceptual similarity of two odorant-mixtures. Therefore, given any two odorant-mixtures, we can predict a significant portion of their ensuing perceptual similarity.

Since food ingredients contain several flavor compounds (Ahn et al. 2011), and dishes contain several ingredients, we can predict the flavor perceptual similarity and dissimilarity of a featured ingredient and a dish to provide quantitative measurement on how the dish enhances or balances the featured ingredient flavor. We describe one approach here and show some results in Table 1, where the personal preference is to enhance the beef flavor of a stew. The formulation of the approach on flavor enhancement is described as below:

$$S_r = \frac{1}{n}\sum_{i=1}^{n} S_i, \text{where } S_i = 100 \times Pr(D > d_i),$$

where the recipe enhancement score (S_r) ranging from 0 to 100 is the average of ingredient scores (S_i) of ingredients in the recipe, and n is the number of ingredients in the recipe. The ingredient score S_i, which is correlated with the angle distance (d_i) of the given ingredient and the featured ingredient beef, is 100 multiplied by the probability of angle distance in food (D) greater than the calculated angle distance (d_i). The flavor compounds constituents of food ingredients can be found in (Ahn et al. 2011), and the aforementioned probability can be calculated from the empirical distribution of paired ingredients angle distances. While the compound concentration in each ingredient should ideally be taken into account, the lack of systematic data prevents us from exploring their impact in this exercise.

Table 1: Enhancement score of beef stew

Ingredient Combination List	Enhancement Score
beef, cabbage, mushroom, potato, mint, sage, bacon, butter	82
beef, mushroom, shellfish, sage, garlic, ginger, butter	64

We comment that there may be other ways to calculate flavor preference score, such as taking the minimum or maximum of the ingredient scores instead of the mean. The goodness of the approach is open for empirical validation. The key idea of using scientific study of human flavor perception for a computational creativity system is a valid step towards building human-level evaluation models.

2.2 Personalized Novelty Assessment

Creativity is only meaningful when there is a human observer, and each observer's world views, culture, life experience, social network are different, so the perception of novelty which is heavily influenced by these factors are inherently different. A parsnip dish may be common to a European consumer, but may be novel to a Chinese consumer. Therefore we need a personalized novelty assessment specific to a targeted observer or a targeted social group.

Bayesian surprise is proposed to quantify the perceived novelty of a newly created artifact (Varshney et al. 2013). The function measures the change in the observer's belief of known artifacts after observing the newly created artifact, where the belief is characterized by the probability distribution of artifacts. The larger the change is the more surprising or more novel the newly created artifact is.

We adopt the use of Bayesian surprise for personalized novelty assessment, and propose to use Internet activity and social media to construct a personalized set of artifacts known to a given individual or a social group. Then, we calculate a personalized surprise score of the newly created artifact. For example, we can learn recipes and ingredients known to an individual from various websites such as *Pinterest* and *allrecipes.com* by gathering recipes posted, reviewed and pined by the individual and her neighborhood in the social network. We denote the frequency of artifact a at time t known to individual p as $f_a(p, t)$. The weighted frequency ($\tilde{f}_a(p, t)$) of artifact known to the individual can be calculated by incorporating social proximity and temporal proximity.

$$\tilde{f}_a(p,t) = \sum_{t'<t} w_T(t',t) \times \Big\{ \sum_{p' \in \text{neighbor of } p} w_S(p',p) \times f_a(p,t) \Big\},$$

where $w_T(t', t)$ and $w_S(p', p)$ are inversely related to temporal proximity and social proximity, respectively. Namely, an artifact which was seen long time ago may be forgotten by the individual (Ebbinghaus 1913), and an artifact known to a closer neighbor in one's social network has higher chance to be known by the individual (Mislove et al. 2007).

Although the ontology to define artifacts and data source may be domain specific, the set forth methodological framework can be extended to other creativity domain for personalized novelty assessment.

3 A Search Method for Recipe Generation

Artifact generation is often a pre-cursor to the evaluation process. A common approach is to rely on human responses to evaluate artifacts such as rhythm and pitch combinations (Monteith, Martinez, and Ventura 2012) and visual narratives (Pérez y Pérez, Morales, and Rodriguez 2012). While this approach is sometimes unavoidable, it is clearly not desirable because it is impossible for humans to explore the entire creativity space and evaluate every newly generated artifact. Recently, there has been a growing interest in the computational creativity community to design evaluation mechanisms that are more robust and objective. (Jordanous 2011) proposed an evaluation guideline for creative systems, (Colton 2008) suggested that *how* a creative work is produced is critical to it being perceived as creative, and (Agustini and Manurung 2012) evaluated the performance of their riddle creation system by comparing the newly created artifacts with those created by another creativity engine. In culinary creativity evaluation, (Morris et al. 2012) trained an artificial neural network model to evaluate the generated recipes, and (Varshney et al. 2013) proposed a cognitive model motivated by the human flavor perception science.

(Boden 1990) proposed that the model of creativity involves a conceptual space and its exploration by creative agents. This conceptual space is a set of artifacts that satisfy certain constraints of the item or idea being generated. (Wiggins 2006) introduced the creative systems framework which revolves around a search mechanism for moving through this conceptual *search* space. For the recipe generation problem, we consider the complexity of creating recipes which may contain 15 or more ingredients. As discussed by (Varshney et al. 2013), the search space for such problems could be in the scale of quintillions (10^{18}) or more. An intelligent search method is necessary to reduce the computational time and guarantee performance. Towards this end, we argue that evaluation should be a part of the generation/search process and propose an optimization-based approach for the recipe creation problem.

The proposed approach models the three evaluation metrics which were discussed in (Varshney et al. 2013) - novelty assessed using Bayesian surprise, flavor pleasantness and food pairing - as the objective function, and the ingredient requirements, identified through learning about the $\langle cusine, dish, ingredient \rangle$ pairing frequencies from the corpus of recipes, as constraints. We are interested in identifying dishes which perform well on all three metrics. The objective function could be formulated in more than one way - maximizing the average of the three metrics, or $max(min(novelty, flavorpleasantness, foodpairing))$. In this paper, we formulate the problem to maximize the average of the three metrics. However, for purposes of performance comparison, we compute the score for individual metrics. The goal is to find a local maxima (or minima) $\mathbf{X}^*$ in terms of the evaluation metrics in the recipe creation space. Let T be a set of ingredient types, I be the set of ingredients, B be set of must have ingredients, C_i be the set of chemical compounds in ingredient i, and R be the set of recipes.

Parameters

p_c	:	pleasantness score of chemical compound c
$\alpha_{i,r}$	:	count of ingredient i in recipe r for a given dish type in the selected cuisine
q^t_{min}	:	minimum quantity of ingredient type t
q^t_{max}	:	maximum quantity of ingredient type t
$P_1(i)$	:	prior belief of ingredient i appearing in a recipe in the selected cuisine
$P_2(i)$	:	posterior belief of ingredient i appearing in a recipe in the selected cuisine

Decision Variables

X_i	:	takes a value of 1 if ingredient i is present in the newly generated recipe and 0 otherwise

$$\max \sum_{i \in I, c \in C_i} X_i p_c + 2 * \frac{\sum_{i,j \in I: i \neq j} X_i X_j \, |C_i \cap C_j|}{\sum_{i \in I} X_i (\sum_{i \in I} X_i - 1)} + \int_I P_2(i) \log \frac{P_2(i)}{P_1(i)}$$

s.t.

$$P_1(i) = \frac{\sum_{r \in R} \alpha_{i,r}}{\sum_{r \in R, i \in I} \alpha_{i,r}} \qquad \forall i \in I \tag{1}$$

$$P_2(i) = \frac{X_i + \sum_{r \in R} \alpha_{i,r}}{\sum_{i \in I} X_i + \sum_{r \in R, i \in I} \alpha_{i,r}} \qquad \forall i \in I \tag{2}$$

$$q^t_{min} \leq \sum_{i \in I \cap t} X_i \leq q^t_{max} \qquad \forall t \in T \tag{3}$$

$$X_b = 1 \qquad \forall b \in B \cap I \tag{4}$$

$$X_i \in \{0, 1\} \qquad \forall i \in I \tag{5}$$

Constraints (1) and (2) define prior and posterior beliefs of an ingredient i appearing in a certain recipe respectively. Constraint (3) enforces the quantity of each ingredient type that the system determines is required to prepare the selected type of dish. Constraint (4) enforces the quantity of user-defined ingredients in the recipe being designed. For example, B could represent user-specifications such as nutritional and/or regional constraints.

The above formulation results in a non-convex, non-linear optimization model with integer variables. Prior works in computational creativity have applied AI-search inspired methods (Wiggins 2006; Morris et al. 2012; Ritchie 2012; Veeramachaneni, Vladislavleva, and O'Reilly 2012) to search problems. In the optimization literature, researchers have used multiple relaxation approaches including branch and bound, Bender's decomposition (You and Grossman 2013), conjugate gradient (Dai and Yuan 1999), interior point methods (Vanderbei and Shanno 1999), and genetic algorithms (Morris et al. 2012). Here, we choose a conjugate gradient approach to solving this model due to its storage, computational and convergence guarantee advantages (Nocedal and Wright 2006). As a first step, we utilize the following inequalities to introduce approximations and convert it into a convex optimization model.

$$2 * \frac{\sum_{i,j \in I: i \neq j} X_i X_j \, |C_i \cap C_j|}{\sum_{i \in I} X_i (\sum_{i \in I} X_i - 1)} \geq 2 * \frac{\sum_{i,j \in I: i \neq j} X_i X_j \gamma_{ij}}{{q^t_{max}}^2} \qquad \forall i, j \in I \cap t \tag{6}$$

$$P_2(i) = \frac{X_i + \sum_{r \in R} \alpha_{i,r}}{\sum_{i \in I} X_i + \sum_{r \in R, i \in I} \alpha_{i,r}} \geq \frac{X_i + \sum_{r \in R} \alpha_{i,r}}{q^t_{max} + \sum_{r \in R, i \in I} \alpha_{i,r}} \qquad \forall i \in I \cap t, r \in R \tag{7}$$

Our solution approach was run on a data set presented in (Varshney et al. 2013), which contains 25,000 recipes available on Wikia. For settings, we chose to prepare a *French soup* containing *beef* as a base ingredient. The conjugate gradient algorithm was made to run for three initial solutions (recipes) and four values of convergence limits. The evaluation metrics were averaged over these 12 runs.

To compare the performance of our algorithm, we also designed recipes, under the same settings, using a sampling approach. The sampling algorithm adds ingredient types such as vegetables, fruits, meat etc. sequentially to the set of existing ingredients, such that the ingredient constraints, represented by constraint (4) are met. Since the search space is in the order of 10^{18}, after each ingredient type is added, it samples a fixed number of recipes that satisfy the constraint set. Then, the final set of sampled recipes are evaluated on the basis on the three metrics. In other words, the sampling approach adopts a generation followed by an evaluation approach. Table 2 summarizes the performance of the conjugate gradient approach compared to the sampling approach.

From the results shown above, we note that the conjugate gradient approach creates higher quality recipes in the conceptual search space, compared to the sampling approach. In particular, it performs better in learning about the non-linear metrics such as novelty and food pairing, and creating recipes that are better in these aspects.

Table 2: Model Results

Problem instance	Size of search space ($x10^{18}$)	Improvement in novelty (%)	Improvement in flavor pleasantness (%)	Improvement in food pairing (%)
1	9,000	67.86	26.83	55.20
2	600	100.00	12.12	43.82
3	8	50.00	6.90	41.26

4 Discussion

In this paper, we report new developments in culinary computational creativity from two aspects: personalization in the evaluation metrics and optimization in the generation process. The idea and framework may be extended to other creative endeavors as well.

We draw inspiration from the science of human flavor perception for personalized flavor preference. Although this is specific to culinary domain, the idea of using principles from scientific study of human, such as psychology, neural and sensory science, may help computational creativity in other domains make progress towards a human level evaluation.

There are vast information on the Internet for us to learn an individual or a targeted social group. Although the ontology to define artifacts and data source may be domain specific, such as the personalized novelty assessment for culinary recipe discussed in this paper, utilizing the Internet to gather personalized information for a computational creativity system is very useful in new product creation where computational creativity can bring business value.

In the recipe generation problem, the creation space is extremely large due to the wide variety of cuisines, dishes and ingredients. We note that such large search spaces are commonly encountered in many other domains (Thornton 2007). Our optimization-based approach has shown superiority over a sampling approach in recipe creation, and it can easily be extended to other creativity endeavors where evaluation metrics are well defined and formulated. Additionally, the generation step could also learn from the changes in the evaluation metrics to prune the space of ingredient combinations. Given the size of the possible artifacts, this would be quite helpful in speeding up the search process and optimizing memory requirements.

References

Agustini, T., and Manurung, R. 2012. Automatic evaluation of punning riddle template extraction. *Proc. Int. Conf. Comput. Creativity*.

Ahn, Y.-Y.; Ahnert, S. E.; Bagrow, J. P.; and Barabási, A.-L. 2011. Flavor network and the principles of food pairing. *SR* 1:196.

Boden, M. 1990. The creative mind. *Abacus*.

Burdock, G. A. 2009. *Fenaroli's Handbook of Flavor Ingredients*. Boca Raton, FL: CRC Press.

Colton, S. 2008. Creativity versus the perception of creativity in computational systems. *Proc. AAAI Spring Symp.*

Dai, Y. H., and Yuan, Y. 1999. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.*

Ebbinghaus, H. 1913. *Memory: A Contribution to Experimental Psychology*. Teachers College, Columbia University.

Haddad, R.; Medhanie, A.; Roth, Y.; Harel, D.; and Sobel, N. 2010. Predicting odor pleasantness with an electronic nose. *PLOSCB* 6(4):e1000740.

Jordanous, A. 2011. Evaluating evaluation: Assessing progress in computational creativity research. *Proc. Int. Conf. Comput. Creativity* 102 – 107.

Khan, R. M.; Luk, C.-H.; Flinker, A.; Aggarwal, A.; Lapid, H.; Haddad, R.; and Sobel, N. 2007. Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world. *JNEU* 27(37):10015–10023.

Mislove, A.; Marcon, M.; Gummadi, K. P.; Druschel, P.; and Bhattacharjee, B. 2007. Measurement and analysis of online social networks. *IMC '07 Proc. ACM SIGCOMM Conf. Internet Measurement* 29–42.

Monteith, K.; Martinez, T.; and Ventura, D. 2012. Automatic generation of melodic accompaniments for lyrics. *Proc. Int. Conf. Comput. Creativity* 87 – 94.

Morris, R. G.; Burton, S. H.; Bodily, P. M.; and Ventura, D. 2012. Soup over bean of pure joy: Culinary ruminations of an artificial chef. *Proc. Int. Conf. Comput. Creativity*.

Nocedal, J., and Wright, S. J. 2006. *Conjugate gradient methods*. Springer, New York.

Pérez y Pérez, R.; Morales, N.; and Rodriguez, L. 2012. Illustrating a computer generated narrative. *Proc. Int. Conf. Comput. Creativity*.

Ritchie, G. 2012. A closer look at creativity as search. *Int. Conf. Comput. Creativity* 41 – 48.

Sawyer, R. K. 2012. *Explaining Creativity: The Science of Human Innovation*. Oxford: Oxford University Press.

Shepherd, G. M. 2006. Smell images and the flavour system in the human brain. *Nature* 444(7117):316–321.

Snitz, K.; Yablonka, A.; Weiss, T.; Frumin, I.; Khan, R. M.; and Sobel, N. 2013. Predicting odor perceptual similarity from odor structure. *PLOSCB* 9(9):e1003184.

Thornton, C. 2007. How thinking inside the box can become thinking outside the box. *Proc. Int. Joint Workshop Comput. Creativity* 113 – 119.

Vanderbei, R. J., and Shanno, D. F. 1999. An interior-point algorithm for nonconvex nonlinear programming. *Comput. Optim. Appli.*

Varshney, L. R.; Pinel, F.; Varshney, K. R.; Bhattacharjya, D.; Schöergendorfer, A.; and Chee, Y. M. 2013. A big data approach to computational creativity. arXiv:1311.1213.

Veeramachaneni, K.; Vladislavleva, E.; and O'Reilly, U.-M. 2012. Knowledge mining sensory evaluation data: genetic programming, statistical techniques, and swarm optimization. *GPEM* 13(1):103–133.

Wiggins, G. A. 2006. Searching for computational creativity. *NGC* 24(3):209–222.

You, F., and Grossman, I. E. 2013. Multicut Benders decomposition algorithm for process supply chain planning under uncertainty. *Ann. Oper. Res* 191 – 211.