Towards Semantic Tea Blend Recommendations

Abstract - Tea enthusiasts may wish to try creating their own tea blends, but be unwilling to potentially waste tea or not have an idea of where to start. We propose a system that can provide recommendations for blends that take into account both brewing temperature and the flavors of the component teas. Our system can easily use new kinds of teas, including tisanes (herbal teas), using reasoning to extract brewing temperature from them, and can also incorporate user feedback on flavor combinations over time using nanopublications.

Introduction/Motivation - Tea is a drink that is consumed worldwide and has potential health benefits. People who drink tea may want to try creating their own blends, but might only have small amounts, or their tea might be expensive, and for someone who has many teas, the number of possible combinations grows quickly.

The wide variety of teas and herbal tisanes available can also be a hurdle to people wanting to explore the field, which is also informed by a variety of world philosophies and traditions. Providing an ontology to describe various steeped drinks - that is, beverages prepared by allowing some plant part to soak in water at or below boiling temperature - helps to disambiguate the various types available, and also allows annotation of different varieties with static information such as recommended brewing temperature. We also provide a backbone for semantic characterization of flavors and flavor compatibilities, which are largely subjective, but can be incorporated into a user-based recommendation system. With objective brewing temperature restrictions and a crowdsourced rating and suggestion framework both tied to a basic hierarchy of teas and tisanes, we provide the semantic structure upon which to build an application that would allow even newcomers to the world of tea to provide relevant feedback on blends they enjoy, as well as to gain recommendations for new blends to try.

There is currently no system for obtaining tea blend recommendations.

Use case - Link
The system allows users to enter teas or tisanes (herbal teas) and uses information about those teas - brewing temperature, derived from type of tea or tisane, along with flavor - to provide different kinds of recommendations. The user can enter multiple teas and receive a recommendation for what combinations would be good to try blending, based on the temperatures that the teas should be brewed at and whether their flavors are considered compatible or not. The user can also enter one tea tea and request suggestions for what to blend with it. Because nanopublications are used to record the sources of flavor compatibility suggestions, users can choose to only consider expert sources or user suggestions for blend recommendations.

Related Work:
There are other ontologies that model plants, such as the USDA National Agricultural Library's Agricultural Thesaurus\(^1\), which we link to and extend when describing, for example, what plant an herbal tea/tisane is made from.

www.nanopub.org defines a nanopublication as “the smallest unit of publishable information: an assertion about anything that can be uniquely identified and attributed to its author”\(^2\). Comprised of three parts - the assertion statement, provenance for that assertion, and publication details about the nanopublication itself - they can be represented in the Resource Description Framework (RDF) alongside other ontology statements to provide both statements and their metadata in a machine-readable format.

The majority of existing literature utilizing nanopublications is centered in the biomedical domain, where the assertions are testable and falsifiable: the provenance of the statement refers to the paper describing the experiments that support the assertion. Golden and Shaw 2015, however, utilize nanopublications in the field of historical period naming, a use case which offers challenges similar to ours. They present nanopublications as assertions and associated provenance demonstrating names and spatial-temporal boundaries for historical eras as definitions well-recognized in the field; this is still of great practical use to historians, in spite of the nature of the encoded knowledge not allowing for experimental testing.

Our work further extends the use of nanopublications to provide concrete machine-readable encodings of discursive information by extending to subjective opinions, as well, regarding both perceived flavor of teas and tisanes, as well as their perceived compatibility with each other.

**Technical Approach:**

[Link to concept map](#)
[Link to ontology](#)

Each tea or tisane has at least two pieces of information associated with it: its flavors, and its brewing temperature. The brewing temperature represents a range of water temperatures ranging from low (160~185°F) to medium (185~205°F) to high (205~212°F). The brewing temperature for a particular tea can be derived from what type of tea it is - green teas have lower brewing temperatures than black teas, for example. For tisanes or herbal teas, the brewing temperature can be derived from the part of the plant that particular tisane is made from. Tisanes made from roots, such as ginger tea, need high brewing temperatures, while those made from leaves or flowers, such as mint tea, can be brewed with a wide range of water temperatures. These specific ranges were chosen because most recommendations about brewing temperatures for different kinds of tea would have the entire recommended range for a particular tea wholly in one of our ranges rather than having that range overlap two of ours (for example, recommending that a certain kind of tea should be brewed with water that is 160~170°F, not 180~190°F).

---

\(^1\) http://agclass.nal.usda.gov/

\(^2\) http://nanopub.org/wordpress/?page_id=65
A tea can also be considered a blend. For example, Earl Grey tea is made from some type of tea (typically black), as well as oil of bergamot. In this case, a blend is defined as consisting of at least one tea and at least one flavoring, such as essential oils or spices.

A blend can also be defined as consisting of at least two teas, and if these two teas meet some criteria, the blend is considered to be a compatible blend. When the user asks for recommendations, these compatible blends are returned. A blend is considered to be compatible if the two teas share at least one temperature range (low, medium, or high) and if their flavors are compatible with each other.

Currently, information as to whether flavors are compatible with each other comes from expert sources, but in the future we would also like to include user feedback to inform flavor compatibilities. Each triple stating that two flavors are compatible with each other acts as the assertion in a nanopublication, which also preserves provenance information. In current usage, nanopublications are utilized primarily in a more academic and scientific sense, where the assertion is supported by some researchers in a study. We propose that nanopublications can also be used to encode subjective information such as flavor compatibility. With an architecture that supports user feedback to construct nanopublications about various combinations of flavors, we would be able to give a relative indication of how much the community likes a particular flavor pairing, and also to allow users to filter compatibilities based on whose preferences they feel match their own.

Discussion
Value of Semantics: There are two major points where semantics helps accomplish this task. The first is that for tisanes or herbal teas, we can reason what the brewing temperature of a tisane is from the plant part used to make that tisane using a property chain. When a tisane has a madeFromPart property linking it to a plant part, such as a berry or root, the property chain connects the brewsAt property of that plant part to the hasBrewingTemperature property of the tisane. This makes it easier to add new tisanes to the ontology, and makes it easier for users to add their own (perhaps home-made) tisanes to be used for recommendations, even if those tisanes are not in the ontology.

For example, mint tea has the madeFromPart property linking it to leafPart. leafPart has the property of brewsAt all three temperature ranges; this is based on recommendations for brewing herbal teas and infusions that suggest that while boiling water is typically recommended, cooler temperatures can also be used for brewing leaves and can result in different flavors.

On the other hand, ginger tea has the madeFromPart property linking it to rootPart, and rootPart has the property of brewsAt only a high temperature. Our sources suggest that roots and other tough parts of the plant - berries and bark - need to be boiled in order for chemicals to infuse into the water.
We investigated connections with the Plant Ontology\(^3\), as well, which is an excellent ontology describing plant development and morphology. However, we ultimately elected not to use it, as our inventory of plant parts are focused on culinary definitions, rather than scientific ones. For example, botanically, an eggplant is considered a berry, while a strawberry is not; as we focus on a culinary domain instead, this distinction is less relevant to our work. At the same time, the plants involved in preparing steeped beverages are important agriculturally because of it, and so connections to agricultural vocabularies are more pertinent.

We also need to only return recommendations where the flavors of the teas in the blend are considered to be compatible. Attempting to establish what flavors are 'compatible' is a difficult task, and we want to use several sources of information for this - both expert chef opinions and user recommendations. We are using nanopublications to capture these, and intend to have a feature that would allow users of the application to filter recommendations if they only want to see those that use expert or user opinions as sources.

So far, we have evaluated our ontology using test cases intended to represent a wide variety of teas and tisanes, both in terms of type and of flavor, based off of real drinks. With these, we can test whether the reasoning works correctly by making sure that steeped drinks with different temperature ranges, or with properties that should lead to different temperature ranges, are not recommended, and check that the flavor compatibilities from expert sources are leading to compatible blends in the reasoner. To test our eventual application, we would use a combination of automatic code tests to ensure it works properly (that the proper recommendations are being received from our test cases, for example, or that queries to external databases are retrieving the correct results) and user testing and feedback to ensure that it is easy to use and provides what features users want.

The nanopublication framework also allows us to capture flavor compatibility from expert sources such as recipe books as well as from site users, along with the relevant provenance information. Encoding the sources alongside each flavor compatibility assertion enables a variety of filtering options for users, such as if they find a subset of other community users whose flavor preferences match their own.

Include a link to your project website: Link

**Future Work:** While currently our system can provide recommendations given information about the teas, in the future, we would like to implement a system that can attempt to retrieve this information, making it easier to use. For example, if it could retrieve the type of the tea, the system can use that to determine the brewing temperature using the temperature range property on that type of tea. It might also be possible to retrieve flavor information. In addition, while we currently have some information about flavor compatibilities from expert (professional chef) opinions, we would also like to build

\(^3\) http://www.plantontology.org/
a module for a tea rating site to capture, record, document, and utilize user feedback in order to inform recommendations.

Our system also has some simplifications that could be addressed in the future. For example, the brewing temperature ranges are a simplification intended to make it easier to give good recommendations and are based on general guidelines. However, specific teas may have different recommended brewing temperatures, and different users may prefer different temperatures. For example, while it is generally suggested that green teas should be brewed at around 170f, because they become bitter at higher temperatures, some people might want a very bitter tea.

Another thing we have not yet tried to address in our system is steeping time. How long one steeps the tea varies considerably from tea to tea, even within the same type, and also varies based on individual taste. Some people prefer stronger tea or weaker tea. A person might enjoy a very strong mint tea but a weaker ginger tea, or stronger jasmine teas but weaker black teas.

For both of these, we could consider using user data from an implemented web application to add to our recommendations and to refine them further based on additional user feedback on these features.

Another feature we could consider adding is recipes. Currently, our system only produces recommendations of the form 'Teas A and B are a compatible blend', but leaves it up to the user to decide if that blend should be a 50/50 mix of the two teas or some other ratio. It might be possible to give more complex recipes in the future.

To further extend the utility of nanopublications for users seeking recommendations, we could also expand the ontology to include relevant user information, such that a user could provide other hints about their preferences, like their favorite commercial blends. A system could then leverage this data to help users find other users with similar likes, or to provide some automatic recommendations for blends to try.

**Conclusion:**
Our ontology is able to provide recommendations for blending teas and tisanes. It does this by taking into account brewing temperature (so as not to recommend a blend consisting of teas with quite different brewing temperatures) and compatibility of the flavors of the teas. In the future, we would like to deploy it as an application that can retrieve information from tea databases and collect user feedback to improve recommendations.

**References**


http://nanopub.org/wordpress/
