Sensemaking around How-to-cook Videos

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

Kyungje Jo
KAIST
School of Computing
kyungjejo@kaist.ac.kr

Seayeon Lee
KAIST
School of Computing
syeonlee@kaist.ac.kr

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

Abstract

We conducted a series of exploratory studies on sensemaking behaviors people exhibit while watching how-to-cook videos. The three different scenarios we examined are a) when people seek for alternatives in ingredients, tools and actions, b) when people seek for explanations or more detail on certain instructions, and c) when people use text search and when people use video when learning how to cook a dish. We found a) people often make arbitrary decisions on substituting ingredients, cooking tools, or cooking actions while following instructions, b) people satisfice by verifying knowledge with little data and not wanting to deviate from the initially chosen video, and c) people use text search for definitions and confirmation of substitutions while they use video search for explanations and precise details for instruction steps.

Author Keywords
How-to videos, sensemaking, video browsing

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; See [http://acm.org/about/class/1998/]: for full list of ACM classifiers. This section is required.

Introduction

Complementary to text recipes, how-to-cook videos are attractive sources for learning because they visually illustrate the tacit, subtle details of instructions. These details help users (i.e. cooks) make judgements and adjust to their own context. For example, if in a video, a fork is used to stir eggs with a certain speed, a user without a fork might use a pair of chopsticks to stir the eggs at a different speed. If in a video, potatoes are fried in a non-coated steel frying pan for 20 minutes with oil, a user might fry for 30 minutes on a coated frying pan with butter, because it seems to "make sense".

However, in order to ensure the intended cooking result, cooking experts point out that these improvisations and substitutions using self-reasoning must be made carefully. There is likely to be more than one reason for using an ingredient, a tool, or even for the speed at which a cooking action must be carried out.

We observed that these reasons behind cooking actions are latent information in how-to instruction steps. For example, the instruction "fry onions until they're brown (i.e. caramelize)" tells us what the direct results of frying onions are, but does not tell us that the taste become sweet and the onion becomes watery, when it is appropriate to fry onions for a sweet taste and when it is not, and how to fry onions and not have them watery, which are all relevant information if a user wants to substitute any of the entities in "fry, onions, until brown".

Exploratory Studies

To assist sensemaking around how-to cook videos, we argue that a video browsing interface must present causal relationships between entities and reasons behind actions. To support browsing and understanding of causal rela-

tionships in how-to-cook videos, we conducted a series of semi-structured interviews and observed how learners make sense of the information presented in how-to-cook videos and how they deal with the contextual discrepancies in ingredients, cooking tools, and cooking actions between the context of the video and that of the user. In the following section, we highlight three user scenarios and sensemaking behaviors we have observed.

Study 1 - Alternative Seeking Scenario

We interviewed five participants who indicated that they cook at least once a week. In this informal study, we asked them to imagine they are baking chocolate chip cookies, and asked them to list ingredients and cooking tools they have at home to engage them in a more personally realistic setting. Then we asked them to freely browse and watch videos to learn how to bake chocolate chip cookies. The participants were encouraged to think aloud.

Sensemaking Behavior 1 - Arbitrary Decision-Making

When asked to cook chocolate chip cookies, participants decided to use white sugar instead of brown sugar. At a glance, such a substitution makes sense. However, the reason for using brown sugar is the stickiness that pulls the dough together, not the sweetness. Hence, using white sugar as a substitute is not appropriate. These arbitrary decision making and self-fixing behaviors happens often when the user references prior knowledge in a different context, i.e. "I used white sugar for making muffins sweet". Transfer of prior knowledge from another context needs better support to inform the user whether the context of the prior knowledge is similar to the present one. After prior knowledge has been transferred successfully, the interface must support updating the user's decision for substitution.

Study 2 - Explanation and Detail Seeking Scenario
We recruited 13 participants who indicated they watch cooking videos at least once a week for a semi-structured interview. We asked our participants when they watch how-to-cook videos and why by replicating their usual cooking video watching experience and by walking us through each stage, involving browsing and watching.

We learned people watch how-to-cook videos either before cooking or while cooking. Before cooking, the participants responded they seek recipe specific characteristics like ingredients, cooking tools, expected cooking time, and the difficulty of cooking. Participants said they decide which video to watch by mentally simulating following e specific video.

While cooking, after having chosen a video to follow, participants responded that they do not seek for extra information outside the video they are watching.

Sensemaking Behavior 2 - Satisficing
There are thousands of how-to videos available for a single dish. Choosing which video to watch and follow is a challenge. Most participants reported that they resort to choosing one of the top three search results.

Even when deciding whether an ingredient or a tool is necessary or not, participants only looked at two or three videos. This small-scale verification can lead to wrong conclusions. Participants indicated that both the video browsing and text recipe browsing interfaces are not scalable enough for them to study evidences for necessities and substitutes. Most participants hesitated deviating from the initially chosen video. If points of confusion arose, they would rather improvise than put extra effort into figuring out the exact answers to the questions they had.

Study 3 - Learning Scenario
Our two previous studies primarily focused on understanding how people watch how-to-cook videos to follow along. We recruited six participants to understand how people watch how-to-cook videos for learning how to cook a dish. We asked our participants to reference how-to materials in both videos and text and generate a set of instructions for a novice friend who needs to learn how to bake spinach pizza, including explanations to any potential questions the friend might have.

Sensemaking Behavior 3 - Choosing Between Video and Text
Participants used text search (Google) for keywords, but used video search (YouTube) to understand how the actions are done. From the participants, we learned that recipe videos are not recipes. They are complementary to each other, rather than substitutes. Recipes are summarizing list of instructions, which lack operational information like which steps can be done in parallel, or which steps are reversible, or when it is ready to go the next step, or what needs to be done to cut an ingredient into the size or shape in the instruction. The questions that the participants anticipated from the novice friend were in four categories: 1) What is this ingredient or a tool? 2) What are my alternatives? 3) Why do I need to do this step? 4) How do I do this step? While questions in the first and second category are mostly asked through text searches, questions in the latter two categories are consulted in how-to videos.

Conclusion
Diverse examples and helping users discover them are important for cooking how-tos. Learning to cook a dish is a complex task that requires continuous sensemaking, and it is special because results of a cooking process are non-binary. It is hard to judge whether someone has successfully followed the instructions or not, nor is it necessary to

follow the instructions exactly. From our exploratory studies, we observed that people make arbitrary decisions, satisfice, and choose consciously between different media of instruction while trying to make sense of the required steps to cook a certain dish. We argue that browsing and watching how-to-cook videos require different strategies than other videos.