Compiling a Lexicon of Cooking Actions for Animation Generation

Kiyoaki Shirai **Hiroshi Ookawa**
Japan Advanced Institute of Science and Technology
1-1, Asahidai, Nomi, 923-1292, Ishikawa, Japan
{kshirai,h-ookawa}@jaist.ac.jp

Abstract

This paper describes a system which generates animations for cooking actions in recipes, to help people understand recipes written in Japanese. The major goal of this research is to increase the scalability of the system, i.e., to develop a system which can handle various kinds of cooking actions. We designed and compiled the lexicon of cooking actions required for the animation generation system. The lexicon includes the action plan used for animation generation, and the information about ingredients upon which the cooking action is taken. Preliminary evaluation shows that our lexicon contains most of the cooking actions that appear in Japanese recipes. We also discuss how to handle linguistic expressions in recipes, which are not included in the lexicon, in order to generate animations for them.

1 Introduction

The ability to visualize procedures or instructions is important for understanding documents that guide or instruct us, such as computer manuals or cooking recipes. We can understand such documents more easily by seeing corresponding figures or animations. Several researchers have studied the visualization of documents (Coyne and Sproat, 2001), including the generation of animation (Andre and Rist, 1996; Towns et al., 1998). Such animation systems help people to understand instructions in documents. Among the various types of documents, this research focuses on the visualization of cooking recipes.

Many studies related to the analysis or generation of cooking recipes have been done (Adachi, 1997; Webber and Eugenio, 1990; Hayashi et al., 2003; Shibata et al., 2003). Especially, several researchers have proposed animation generation systems in the cooking domain. Karlin, for example, developed SEAFACT (Semantic Analysis For the Animation of Cooking Tasks), which analyzed verbal modifiers to determine several features of an action, such as the aspectual category of an event, the number of repetitions, duration, speed, and so on (Karlin, 1988). Uematsu developed "Captain Cook," which generated animations from cooking recipes written in Japanese (Uematsu et al., 2001). However, these previous works did not mention the scalability of the systems. There are many linguistic expressions in the cooking domain, but it is uncertain to what extent these systems can convert them to animations.

This paper also aims at developing a system to generate animations from cooking recipes written in Japanese. We especially focused on increasing the variety of recipes that could be accepted. After presenting an overview of our proposed system in Subsections 2.1 and 2.2, the more concrete goals of this paper will be described in Subsection 2.3.

2 Proposed System

2.1 Overview

The overview of our animation generation system is as follows. The system displays a cooking recipe in a browser. As in a typical recipe, cooking instructions are displayed step by step, and sentences or phrases representing a cooking action in the recipe are highlighted. When a user does not understand a certain cooking action, he/she can click the highlighted sentence/phrase. Then the system will show the corresponding animation to help the user understand the cooking instruction.

Note that the system does not show all procedures in a recipe like a movie, but generates an animation of a single action on demand. Furthermore, we do not aim at the reproduction of recipe sentences in detail. Especially, we will not prepare object data for many different kinds of ingredients. For example, suppose that the system has object data for a mackerel, but not for a sardine. When a user clicks the sentence "fillet a sardine" to see the animation, the system will show how to fillet a "mackerel" instead of "sardine", with a note indicating that the ingredient is different. We believe

Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, pages 771–778,
Sydney, July 2006. ©2006 Association for Computational Linguistics

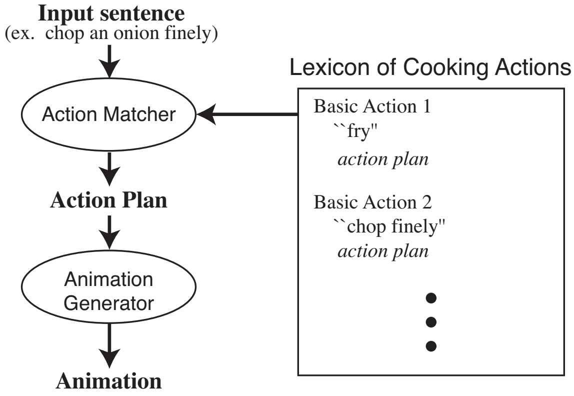

Figure 1: System Architecture

that the user will be more interested in "how to fillet" than in the specific ingredient to be filleted. In other words, the animation of the action will be equally helpful as long as the ingredients are similar. Thus we will not make a great effort to prepare animations for many kinds of ingredients. Instead, we will focus on producing the various kinds of cooking actions, to support users in understanding cooking instructions in recipes.

2.2 System Architecture

Figure 1 illustrates the architecture of the proposed system. First, we prepare the lexicon of cooking actions. This is the collection of cooking actions such as "fry", "chop finely", etc. The lexicon has enough knowledge to generate an animation for each cooking action. Figure 2 shows an example of an entry in the lexicon. In the figure, "*expression*" is a linguistic expression for the action; "*action plan*" is a sequence of action primitives, which are the minimum action units for animation generation. Roughly speaking, the action plan in Figure 2 represents a series of primitive actions, such as cutting and rotating an ingredient, for the basic action "chop finely". The system will generate an animation according to the action plan in the lexicon. Other features, "*ingredient examples*" and "*ingredient requirement*", will be explained later.

The process of generating an animation is as follows. First, as shown in Figure 1, the system compares an input sentence and *expression* of the entries in the lexicon of cooking actions, and finds the appropriate cooking action. This is done by the module "*Action Matcher*". Then, the system extracts an action plan from the lexicon and passes it to the "*Animation Generator*" module. Finally *Animation Generator* interprets the action plan and produces the animation.

2.3 Goal

The major goals of this paper are summarized as follows:

G1. Construct a large-scale lexicon of cooking actions

In order to generate animations for various kinds of cooking actions, we must prepare a lexicon containing many basic actions.

G2. Handle a variety of linguistic expressions

Various linguistic expressions for cooking actions may occur in recipes. It is not realistic to include all possible expressions in the lexicon. Therefore, when a linguistic expression in an input sentence is not included in the lexicon, the system should calculate the similarity between it and the basic action in the lexicon, and find an equivalent or almost similar action.

G3. Include information about acceptable ingredients in the lexicon

Even though linguistic expressions are the same, cooking actions may be different according to the ingredient upon which the action is taken. For example, "cut into fine strips" may stand for several different cooking actions. That is, the action of "cut **cucumber** into fine strips" may be different than "cut **cabbage** into fine strips", because the shapes of cucumber and cabbage are rather different. Therefore, each entry in the lexicon should include information about what kinds of ingredients are acceptable for a certain cooking action.

As mentioned earlier, the main goal of this research is to increase the scalability of the system, i.e., to develop an animation generation system that can handle various cooking actions. We hope that this can be accomplished through goals G1 and G2.

In the rest of this paper, Section 3 describes how to define the set of actions to be compiled into the lexicon of cooking actions. This concerns goal G1. Section 4 explains two major features in the lexicon, "*action plan*" and "*ingredient requirement*". The feature *ingredient requirement* is

Basic Action 2	
expression	みじん切りにする (chop finely)

action plan	cut(*ingredient*,*utensil*,*location*, 2)
	rotate(*ingredient*,*location*, *x*, 90)
	cut(*ingredient*,*utensil*,*location*,20)
	rotate(*ingredient*,*location*, *z*, 90)
	cut2(*ingredient*,*utensil*,*location*, 10)
	cut(*ingredient*,*utensil*,*location*, 20)

ingredient examples	おくら (okra), しいたけ (shiitake mushroom)
ingredient requirement	kind=vegetable|mushroom

Figure 2: Example of an Entry in the Lexicon of Cooking Actions

related to goal G3. Section 5 reports a preliminary survey to construct the module *Action Matcher* in Figure 1, which is related to goal G2. Finally, Section 6 concludes the paper.

3 Defining the Set of Basic Actions

In this and the following sections, we will explain how to construct the lexicon of cooking actions. The first step in constructing the lexicon is to define the set of basic actions. As mentioned earlier (goal G1 in Subsection 2.3), a large-scale lexicon is required for our system. Therefore, the set of basic actions should include various kinds of cooking actions.

3.1 Procedure

We referred to three cooking textbooks or manuals (Atsuta, 2004; Fujino, 2003; Takashiro and Kenmizaki, 2004) in Japanese to define the set of basic actions. These books explain the fundamental cooking operations with pictures, e.g., how to cut, roast, or remove skins/seeds for various kinds of ingredients. We extracted the cooking operations explained in these three textbooks, and defined them as the basic actions for the lexicon. In other words, we defined the basic actions according to the cooking textbooks. The reasons why we used the cooking manuals as the standard for the basic actions are summarized as follows:

1. The aim of cooking manuals used here is to comprehensively explain basic cooking operations. Therefore, we expect that we can collect an exhaustive set of basic actions in the cooking domain.

2. Cooking manuals are for beginners. The aim of animation generation system is to help people, especially novices, to understand cooking actions in recipes. The lexicon of cooking actions based on the cooking textbooks includes many cooking operations that novices may not know well.

3. The definition of basic actions does not depend on the module *Animation Generator*.

   One of the standards for the definition of basic actions is animations generated by the system. That is, we can define basic cooking actions so that each cooking action corresponds to an unique animation. This approach seems to be reasonable for an animation generation system; however, it depends on the module *Animation Generator* in Figure 1. Many kinds of rendering engines are now available to generate animations. Therefore, *Animation Generator* can be implemented in various ways. When changing the rendering engine used in *Animation Generator*, the lexicon of cooking actions must also be changed. So we decided that it would not be desirable to define the set of basic actions according to their corresponding animations.

   In our framework, the definition of basic actions in the lexicon does not depend on *Animation Generator*. This enables us to use any kind of rendering engine to produce an animation. For example, when we use a poor engine and want to design the system so that it generates the same animation for two or more basic actions, we just describe the same *action plan* for these actions.

We manually excerpted 267 basic actions from three cooking textbooks. Although it is just a collection of basic actions, we refer it as the initial

Table 1: Examples of Basic Actions

expression	ingredient examples
三枚におろす (fillet)	あじ (mackerel)
炊き込む (boil)	
炊く (boil)	
くし形切りにする (cut into a comb shape)	トマト (tomato), じゃがいも (potato)
くし形切りにする (cut into a comb shape)	かぼちゃ (pumpkin)
くし形切りにする (cut into a comb shape)	カブ (turnip)

lexicon of cooking actions. Table 1 illustrates several examples of basic actions in the initial lexicon. In the cooking manuals, every cooking operation is illustrated with pictures. "*Ingredient examples*" indicates ingredients in pictures used to explain cooking actions.

3.2 Preliminary Evaluation

A preliminary experiment was conducted to evaluate the scalability of our initial lexicon of basic actions. The aim of this experiment was to check how many cooking actions appearing in real recipes are included in the initial lexicon.

First, we collected 200 recipes which are available on web pages [1]. We refer to this recipe corpus as R_a hereafter. Next, we analyzed the sentences in R_a and automatically extracted verbal phrases representing cooking actions. We used JUMAN [2] for word segmentation and part-of-speech tagging, and KNP [3] for syntactic analysis. Finally, we manually checked whether each extracted verbal phrase could be matched to one of the basic actions in the initial lexicon.

Table 2 (A) shows the result of our survey. The number of basic actions was 267 (a). Among these actions, 145 (54.3%) actions occurred in R_a (a1). About half of the actions in the initial lexicon did not occur in the recipe corpus. We guessed that this was because the size of the recipe corpus was not very large.

The number of verbal phrases in R_a was 3977 (b). We classified them into the following five cases: (b1) the verbal phrase corresponded with one of the basic actions in the initial lexicon, and its linguistic expression was the same as one in the lexicon; (b2) the verbal phrase corresponded with a basic action, but its linguistic expression differed from one in the lexicon; (b3) no corresponding basic action was found in the initial lexicon, (b4) the extracted phrase was not a verbal phrase, caused by error in analysis, (b5) the verbal phrase did not stand for a cooking action. Note that the cases in which verbal phrases should be converted to animations were (b1), (b2) and (b3). The numbers in parentheses (...) indicate the ratio of each case to the total number of verbal phrases, while numbers in square brackets [...] indicate a ratio of each case to the total number of (b1), (b2) and (b3).

We expected that the verbal phrases in (b1) and (b2) could be handled by our animation generation system because the initial lexicon contained the corresponding basic actions. On the other hand, our system cannot generate animations for verbal phrases in (b3), which was 42.3% of the verbal phrases our system should handle. Thus the applicability of the initial lexicon was poor.

3.3 Adding Basic Actions from Recipe Corpus

We have examined what kinds of verbal phrases were in (b3). We found that there were many general verbs, such as "加える (add)", "入れる (put in)", "熱する (heat)", "付ける (attach)", "のせる (put on)", etc. Such general actions were not included in the initial lexicon, because we constructed it by extracting basic actions from cooking textbooks, and such general actions are not explained in these books.

In order to increase the scalability of the lexicon of cooking actions, we selected verbs satisfying the following conditions: (1) no corresponding basic action was found in the lexicon for a verb; (2) a verb occurred more than 10 times in R_a. In all, 31 verbs were found and added to the lexicon as new basic actions. It is undesirable to define basic actions in this way, because the lexicon may then depend on a particular recipe corpus. However, we believe that the new basic actions are very general, and can be regarded as almost independent of with the corpus from which they were extracted.

In order to evaluate the new lexicon, we prepared another 50 cooking recipes (R_b hereafter). Then we classified the verbal phrases in R_b in the same way as in Subsection 3.2. The results are shown in Table 2 (B). Notice that the ratio

[1] http://www.bob-an.com/

[2] http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html

[3] http://www.kc.t.u-tokyo.ac.jp/nl-resource/knp.html

Table 2: Result of Preliminary Evaluation

(A) Survey on R_a

(a)	# of basic actions	267
(a1)	basic actions occurred in R_a	145 (54.3%)
(b)	# of verbal phrases	3977
(b1)	basic action(same)	974 (24.5%) [28.0%]
(b2)	basic action(dif.)	1031 (25.9%) [29.7%]
(b3)	not basic action	1469 (36.9%) [42.3%]
(b4)	analysis error	180 (4.5%)
(b5)	not cooking action	323 (8.1%)

(B) Survey on R_b

(a)	298
(a1)	106 (35.6%)
(b)	959
(b1)	521 (54.3%) [62.2%]
(b2)	262 (27.3%) [31.3%]
(b3)	55 (5.7%) [6.6%]
(b4)	45 (4.7%)
(b5)	76 (7.9%)

of the number of verbal phrases contained in the lexicon to the total number of target verb phrases was 94.5% ((b1)62.2% + (b2)31.3%). This is much greater than the ratio in Table 2 (A) (57.7%). Therefore, although the size of test corpus is small, we hope that the scalability of our lexicon is large enough to generate animations for most of the verbal phrases in cooking recipes.

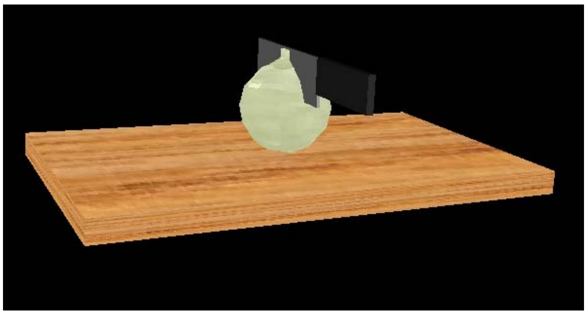

Figure 3: Snapshot of Generated Animation

4 Compilation of the Lexicon of Basic Actions

After defining the set of basic actions for the lexicon, the information of each basic action must be described. As shown in Figure 2, the main features in our lexicon are *expression*, *action plan*, *ingredient examples* and *ingredient requirement*. The term *expression* stands for linguistic expressions of basic actions, while *ingredient examples* stands for examples of ingredients described in the cooking manuals we referred to when defining the set of basic actions. As shown in Table 1, these two features have already been included in the initial lexicon created by the procedure in Section 3. This section describes the compilation of the rest of the features: *action plan* in Subsection 4.1 and *ingredient requirement* in Subsection 4.2.

4.1 Action Plan

For each basic action in the lexicon, the action plan to generate the corresponding animation is described. *Action plan* is the sequence of action primitives as shown in Figure 2. Of the 298 basic actions in the lexicon, we have currently described action plans for only 80 actions. Most of them are actions to cut something.

We have also started to develop *Animation Generator* (see Figure 1), which is the module that interprets action plans and generates animations. We used VRML for animation generation. Figure 3 is a snapshot of the animation for the basic action "みじん切りにする (chop finely)" generated by our system.

Our current focus has been on the design and development of the lexicon of cooking actions, rather than on animation generation. Implementation of the complete *Animation Generator* as well as a description of the action plans for all basic actions in the lexicon are important future works.

4.2 Ingredient Requirement

Several basic actions have the same expression in our lexicon. For instance, in Figure 1, there are three basic actions represented by the same linguistic expression "くし形切りにする (cut into a comb shape)". These three actions stand for different cooking actions. The first one stands for the action used to cut something like a "tomato" or "potato" into a comb shape. The second stands for the following sequence of actions: first cut something in half, remove its core or seeds, and cut it into a comb shape. This action is taken on pumpkin, for instance. The third action represents the cooking action for "turnip": remove the leaves of the turnip and cut it into a comb shape. In other words, there are different ways to cut different ingredients into a comb shape. Differences among these actions depend on what kinds of ingredients are to be cut.

As described in Section 2.2, the module *Action Matcher* accepts a sentence or phrase for which a user wants to see the animation, then finds a corresponding basic action from the lexicon. In order to find an appropriate basic action for a recipe sentence, the lexicon of cooking actions should include information about what kinds of ingredients are acceptable for each basic action. Note that the judgment as to whether an ingredient is suitable or not highly depends on its features such as kind, shape, and components (seed, peel etc.) of the ingredient. Therefore, the lexicon should include information about what features of the ingredients must be operated upon by the basic actions.

For the above reason, *ingredient requirement* was introduced in the lexicon of cooking actions. In this field, we manually describe the required features of ingredients for each basic action. Figure 4 illustrates the three basic actions of くし形切りにする (chop into a comb shape) in the lexicon [4]. The basic action *a1*, "kind=vegetable, shape=sphere" in *ingredient requirement*, means that only a vegetable whose shape is spherical is acceptable as an ingredient for this cooking action. On the other hand, for the basic action *a2*, only a vegetable whose shape is spherical and containing seeds is acceptable. For *a3*, "instance=カブ (turnip)" means that only a turnip is suitable for this action. In our lexicon, such specific cooking actions are also included when the reference cookbooks illustrate special cooking actions for certain ingredients. In this case, a cookbook illustrates cutting a *turnip* into a comb shape in a different way than for other ingredients.

4.2.1 Feature Set of Ingredient Requirement

Here are all the attributes and possible values prepared for the *ingredient requirement* field:

- kind

  This attribute specifies kinds of ingredients. The possible values are:

  vegetable, mushroom, fruit, meat,
  fish, shellfish, seafood, condiment

  "Seafood" means seafood other than fish or shellfish, such as イカ (squid), タラコ (cod roe) and so on.

- veg

  This attribute specifies subtypes of vegetables. Possible values for this attribute are "green", "root" and "layer". "Green" stands for green vegetables such as ほうれん草 (spinach) and 白菜 (Chinese cabbage). "Root" stands for root vegetables such as じゃがいも (potato) and ごぼう (burdock). "Layer" stands for vegetables consisting of layers of edible leaves such as レタス (lettuce) and キャベツ (cabbage).

- shape

  This attribute specifies shapes of ingredients. The possible values are:

  sphere, stick, cube, oval, plate, filiform

- peel, seed, core

  These attributes specify various components of ingredients. Values are always 1. For example, "peel=1" stands for ingredients with peel.

- instance

  This specifies a certain ingredient, as shown in basic action *a3* in Figure 4.

The information about ingredient requirements was added for 186 basic actions out of the 298 actions in the lexicon. No requirement was needed for the other actions, i.e., these actions accept any kind of ingredients.

4.2.2 Lexicon of Ingredients

In addition to the lexicon of cooking actions, the lexicon of ingredients is also required for our system. It includes ingredients and their features such as kind, shape and components. We believe that this is domain-specific knowledge for the cooking domain. Thesauri or other general-purpose language resources would not provide such information. Therefore, we newly compiled the lexicon of ingredients, which consists of only those ingredients appearing in the *ingredients example* in the lexicon of cooking actions. The number of ingredients included in the lexicon is 93. For each entry, features of the ingredient are described. The feature set used for this lexicon is the same as that for the *ingredient requirement* described in 4.2.1, except for the "instance" attribute.

[4] *action plan* is omitted in Figure 4.

Basic Action *a1*	
expression	くし形切りにする (cut into a comb shape)
ingredient examples	トマト (tomato), じゃがいも (potato)
ingredient requirement	kind=vegetable, shape=sphere
Basic Action *a2*	
expression	くし形切りにする (cut into a comb shape)
ingredient examples	かぼちゃ (pumpkin)
ingredient requirement	kind=vegetable, shape=sphere, seed=1
Basic Action *a3*	
expression	くし形切りにする (cut into a comb shape)
ingredient examples	カブ (turnip)
ingredient requirement	instance=カブ (turnip)

Figure 4: Three Basic Actions of "くし形切りにする (cut into a comb shape)"

The current lexicon of ingredients is too small. Only 93 ingredients are included. A larger lexicon is required to handle various recipe sentences. In order to enlarge the lexicon of ingredients, we will investigate a method for the automatically acquisition of new ingredients with their features from a collection of recipe documents.

5 Matching between Actions in a Recipe and the Lexicon

Action Matcher in Figure 1 is the module which accepts a recipe sentence and finds a basic action corresponding to it from the lexicon. One of the biggest difficulties in developing this module is that linguistic expressions in a recipe may differ from those in the lexicon. So we have to consider a flexible matching algorithm between them.

To construct *Action Matcher*, we refer to the verbal phrases classified in (b2) in Table 2. Note that the linguistic expressions of these verbal phrases are inconsistent with the expressions in the lexicon. We examined the major causes of inconsistency for these verbal phrases. In this paper, we will report the result of our analysis, and suggest some possible ways to find the equivalent action even when the linguistic expressions in a recipe and the lexicon are different. The realization of *Action Matcher* still remains as future work.

Figure 5 shows some examples of observed inconsistency in linguistic expressions. In Figure 5, the left hand side represents verbal phrases in recipes, while the right hand side represents expressions in the lexicon of cooking actions. A slash indicates word segmentation. Causes of inconsistency in linguistic expressions are classified as follows:

- Inconsistency in word segmentation

  Word segmentation of verbal phrases in recipes, as automatically given by a morphological analyzer, is different from one of the basic actions in the lexicon, as shown in Figure 5 (a).

  In order to succeed in matching, we need an operation to concatenate two or more morphemes in a phrase or to divide a morpheme into to two or more, then try to check the equivalence of both expressions.

- Inconsistency in case fillers

  Verbs in a recipe and the lexicon agree, but their case fillers are different. For instance, in Figure 5 (b), the verb "ふる (sprinkle)" is the same, but the accusative case fillers "唐辛子 (chili)" and "塩 (salt)" are different. In this case, we can regard both as representing the same action: to sprinkle a kind of condiment.

  In this case, the lexicon of ingredients (see 4.2.2) would be helpful for matching. That is, if both 唐辛子 (chili) and 塩 (salt) have the same feature "kind=condiment" in the lexicon of ingredients, we can judge that the phrase "唐辛子/を/ふる (sprinkle chili)" corresponds to the basic action "塩/を/ふる (sprinkle salt)".

- Inconsistency in verbs

  Disagreement between verbs in a recipe and the lexicon is one of the major causes of inconsistency. See Figure 5 (c), for instance.

	Expressions in Recipes	Expressions in Lexicon
(a)	割り / ほぐす · · ·break (egg) (divide) (loosen)	割りほぐす · · ·break (egg) (break)
(b)	唐辛子 / を / ふる · · ·sprinkle chili (chili) (ACC) (sprinkle)	塩 / を / ふる · · ·sprinkle salt (salt) (ACC) (sprinkle)
(c)	砂出し / を /する · · ·make (shellfish) (Spewing sand) (ACC) (do) spew out sand	塩水 / に /ひたす · · ·dip it into (salt water) (LOC) (dip) salt water

Figure 5: Inconsistency in Linguistic Expressions

These two phrases represent the same action [5], but the linguistic expressions are totally different.

In this case, the matching between them is rather difficult. One solution would be to describe all equivalent expressions for each action in the lexicon. Since it is not realistic to list equivalent expressions exhaustively, however, we want to automatically collect pairs of equivalent expressions from a large recipe corpus.

6 Conclusion

In this paper, we have described the basic idea for a system to generate animations for cooking actions in recipes. Although the system is not yet complete and much work still remains to be done, the main contribution of this paper is to show the direction for improving the scalability of the system. First, we designed a lexicon of cooking actions including information about action plans and ingredient requirements, which are needed to generate the appropriate cooking animations. We also showed that our lexicon covers most of the cooking actions appearing in recipes. Furthermore, we analyzed the recipe corpus and investigated how to match actions in a recipe to the corresponding basic action in the lexicon, even when they have different linguistic expressions. Such a flexible matching method would also increase the scalability of the system.

References

Hisahiro Adachi. 1997. GCD: A generation method of cooking definitions based on similarity between a couple of recipes. In *Proceedings of the Natural Language Processing Pacific Rim Symposium*, pages 135–140.

Elisabeth Andre and Thomas Rist. 1996. Coping with temporal constraints in multimedia presentation planning. In *Proceedings of the National Conference on Artificial Intelligence*, pages 142–147.

Yoko Atsuta. 2004. *How to cut vegetables (in Japanese)*. Syûeisha.

Bob Coyne and Richard Sproat. 2001. WordsEye: An automatic text-to-scene conversion system. In *Proceedings of the SIGGRAPH*, pages 487–496.

Yoshiko Fujino. 2003. *New Fundamental Cooking (in Japanese)*. SS Communications.

Eri Hayashi, Suguru Yoshioka, and Satoshi Tojo. 2003. Automatic generation of event structure for Japanese cooking recipes (in Japanese). *Journal of Natural Language Processing*, 10(2):3–17.

Robin F. Karlin. 1988. Defining the semantics of verbal modifiers in the domain of cooking tasks. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 61–67.

Tomohide Shibata, Daisuke Kawahara, Masashi Okamoto, Sadao Kurohashi, and Toyoaki Nishida. 2003. Structural analysis of instruction utterances. In *Proceedings of the Seventh International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES2003)*, pages 1054–1061.

Junko Takashiro and Satomi Kenmizaki. 2004. *Standard Cooking: Fundamentals of Cooking (in Japanese)*. Shôgakukan.

Stuart G. Towns, Charles B. Callaway, and James C. Lester. 1998. Generating coordinated natural language and 3D animations for complex spatial explanations. In *Proceedings of the National Conference on Artificial Intelligence*, pages 112–119.

Hideki Uematsu, Akira Shimazu, and Manabu Okumura. 2001. Generation of 3D CG animations from recipe sentences. In *Proceedings of the Natural Language Processing Pacific Rim Symposium*, pages 461–466.

Bonnie Lynn Webber and Barbara Di Eugenio. 1990. Free adjuncts in natural language instructions. In *Proceedings of the International Conference on Computational Linguistics*, pages 395–400.

[5] Note that it is required to dip shellfish into salt water in order to make it spew out sand.