MuST: A Workshop on Multimodal Summarization for Trend Information

Tsuneaki Kato
The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan
kato@boz.c.u-tokyo.ac.jp

Mitsunori Matsushita
Nippon Telegraph and Telephone Corp. NTT Communication Science Labs.
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan
mat@cslab.kecl.ntt.co.jp

Noriko Kando
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
kando@nii.ac.jp

Abstract

The MuST (Multimodal Summarization for Trend Information) workshop was designed to encourage cooperative and competitive studies on summarization and visualization for trend information. The main objective of the workshop is to develop technologies using multimedia presentation that will allow intelligent systems to provide users with appropriate answers to their queries on trend information. These technologies not only could be considered as multimedia presentation generation and multimodal dialogue processing, but also rely largely on information access technologies such as automatic summarization and information extraction, and information visualization. Therefore, the workshop is expected to encourage studies in a wide variety of research fields. A noteworthy feature of the workshop is that the participants share the same research resource, whereby they address common or related themes, with the expectation of encouraging active research and discussion, conforming communities, and constructing and accumulating resources such as tools and corpora.

1 Introduction

The MuST workshop on summarization and visualization for trend information was designed to encourage cooperative and competitive studies in this field by promoting discussion, conforming communities, and constructing and accumulating resources such as tools and corpora. At present, objective and quantitative evaluation using a test set is outside the scope of the workshop. This characteristic makes MuST incomparable to other tasks in the NTCIR workshop. In this paper, we describe the subjects of MuST and their importance, and explain the data set, which plays an important role as a shared resource and creates the centripetal force of the workshop. Then, we state the course, current situation, and future plans for the workshop.

2 The Subjects: Trend Information

Trend information is basically a kind of summarization of temporal statistical data, obtained through synthesis rather than simple enumeration. Changes in product price and sales, the financial status of a company, and the public approval rating of a cabinet or political party are examples of trend information. This is not always single-dimensional temporal information; it can also be multidimensional, including other axes such as agents and locations. For example, the trend in the market share of a given product should be discussed with reference to many competitive companies, and the trend in land price should be depicted as a function of time and location. Many attaches considerable importance to various trends, i.e. how the price of gasoline shifted during the year, what the situation has been over the last few years in the domestic personal computer market, how Sony has been doing recently, how terrible the typhoons were last autumn, and what the home run tallies are in major league base-

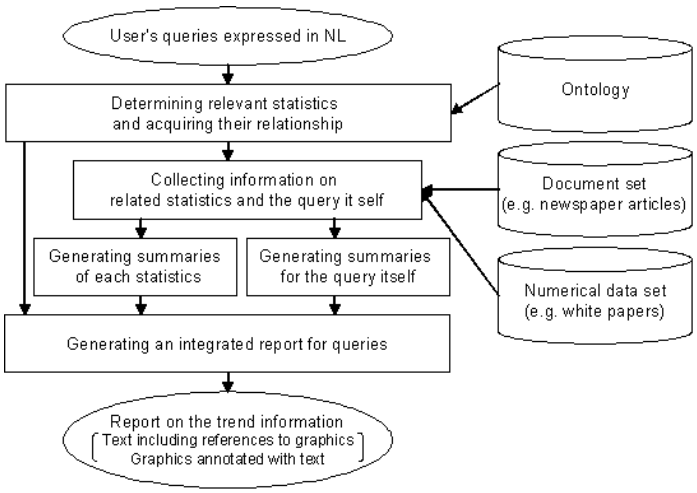

Figure 1. Framework of summarization and visualization for trend information

ball this season. The subject of MuST is the development of systems and technologies that provide users with answers to those queries on trends using concise, plain text and information graphics, or using multimedia presentation, in which the two media work in a cooperative manner.

Figure 1 shows the proposed framework of summarization and visualization for trend information. A user's query on the trend of the personal computer industry in 1998, for example, is processed as follows. Using some type of ontology, the system determines that statistics such as shipment volume, shipment value and market share of major makers are needed for describing that trend. Then, it also understands the relationship among the statistics. Related articles and/or data on those statistics are retrieved from a background document set such as newspaper articles and/or a background numerical data set such as white papers. In the case of newspaper articles, since only certain portions of each retrieved article are likely to be relevant, the system must extract those portions and categorize them according to their relevant. The portions to be extracted include those related to the user's query itself, such as "the personal computer market, which had been growing steadily, now suffers from depression and saturated growth," and those mentioning each related statistic, such as "the domestic shipment volume for the first half of the year was 4,391,000, which is 34% higher than in the same period last year and marked the highest record for the half-year range." From the sets of extracted portions, the system generates summaries. Summarization is conducted by each statistics and the user's query itself. The summaries obtained are not restricted to text format, but can also be graphics such as charts visualizing the change in given statistics. In such a case, the summarization process may look like information extraction [11]. Then, the numerical data collected is referred to and incorporated into the summaries. We are convinced that textual information plays a more important role than numerical data in generating trend information, since evaluation and judgment is crucial and cannot be obtained from pure numerical data. Finally, the summaries are integrated into one report with consideration given to the relationship between statistics, such as "when unit prices become lower, the growth in shipment values does not increase even when shipment volumes increase." The report can take several different styles: from text or graphics alone, to combinations such as text that includes references to graphics or graphics annotated with text.

Within this framework, the first part, determining the relevant statistics and understanding their relationship, concerns knowledge engineering, which addresses the matter of how to acquire and organize knowledge on what statistics are related to a given query; what relationship exists among them and so on. We will have to wait for further progress in ontology engineering and semantic Web technologies for this subject. The second part, collecting the information, has very few problems and can be achieved using the current information retrieval technologies. The remaining components are the subjects of MuST: Summarizing as text and/or information graphics the information on a specific topic, especially on a given statistic, distributed in multiple texts and/or numerical data such as tables, and generating reports or multimedia presentation by integrating and organizing those summaries in order to answer the user's query.

3 Importance of the Subjects

The subjects of MuST are related to several fields of research and have their own place of importance in those fields [4]. First, they can be considered as multimedia presentation generation and multimodal dia-

logue processing if we emphasize how the technologies allow systems to provide users with appropriate answers to their queries using multimedia presentation. From this perspective, MuST includes novel challenges to content understanding for multimedia presentation generation, since the utilization of text extracted from newspaper articles as the knowledge source requires that we obtain formal representation of the content to be presented or we must explore presentation generation without formal representation of the content. This is in sharp contrast to the traditional approach because it is assumed that formal representation is given.

Second, focusing on summarizing and generating trend information from texts, we realize that these subjects are related to information access technologies such as question answering, automatic summarization, and information extraction. For question answering, the subjects broaden the current framework, the main concern of which is answering factoid questions [10][1], and move toward integration with automatic summarization, since it is not sufficient to extract simple facts; what is needed is to collect information from multiple documents and compile the information into an answer to a given query. For automatic summarization, MuST raises the problem of user-focused multi-document summarization, which is a major concern in the latest research [6], since the material to be summarized is scattered through several articles and the summaries expected are not summaries of the whole of those articles but instead are those of only the portions relevant to the user's query. In such multi-document summarization, it is important to detect duplications and reduce redundancy [2]. Material for trend information is quite redundant since frequent mention is made of previous values in comparison with current ones and must be an excellent test bed for those technologies. Because of handling facts related to numerical values, these subjects also need domain-independent information extraction for several types of statistics that appear related to the trend information. Generalization and the pursuit of domain independency are the current topics in those technologies [7], and MuST promotes efforts in this direction.

Third, visualization of trend information, that is, visualizing temporal changes of statistical values and their relationships in an intuitive and easy-to-understand manner, can be a foundation of visualization since it is necessary to establish a visualization technology applicable for a wide range of statistics. On the other hand, if the data from such visualization comes as a result of information extraction from texts, some inference is required to compensate for the sparseness of data obtained. This is in contrast to conventional visualization, which is used for making an excessive amount of data intuitively understandable, and needs a novel framework for visualization. In addition, one of the major objectives of information visualization is information outlining, in which a huge amount of information is made easy accessible through its arrangement in visually understandable spaces [9][8]. This is also the case in visualization for trend information. That is, multimedia presentation generated by those technologies can be used as a method for accessing documents containing the original information. From this point of view also, MuST provides attractive approaches to information visualization.

Moreover, these subjects are related to several elemental technologies on natural language understanding and soft computing. For example, it is important to investigate the relationship between the vague expressions of degree in language and quantitative numerical changes, and the linguistic roles of tense and aspect for expressing changes. Addressing not only the construction of application systems but also their many elemental technologies is very important.

4 The Framework

In order to encourage these studies, MuST participants share the same research resource, whereby they can address common or related themes. The shared resource plays an important role in motivating participation in the workshop and creates the centripetal force of the workshop. It should be a substitute for objective evaluation using a common test set in evaluation workshops. The shared resource of MuST includes not only the materials to be processed but also the intermediate results of the process, which behave as a hub in the sense that it is the output of some modules and also the input of other modules within the framework proposed as the research model, as previously explained. In addition, sample summaries and information graphics are also included in the data set as a reference for the output design. Using this data set with the research model makes it possible for the participants to address their own subjects that constitute some part of the whole framework. In some research, the data set is referred to as a model for the output, while in other research it is used as sample input. Participants can discuss their parts with each other using the data set as common ground, and can also position their studies or their modules developed within the framework. This is especially important for those studying elemental technologies.

Providing this data set is the primary role of the workshop organizers, who also provide the means for holding discussions by setting up a mailing list and securing opportunities for research presentations, which also encourages research. The participants are requested to present the results and/or progress of their research using this data set at an occasion specified by the organizers. The presentation is made for a closed

audience, that is, the participants of MuST, and does not prevent the participants from making the same presentation on other occasions. Making the presentation and joining the mailing list are the only obligations of the participants; providing tools or corpora is done on a voluntary basis.

5 The Data Set

The MuST data set is designed as the centripetal force of the workshop for accelerating research on summarization and visualization for trend information. The design principle and specifications are as follows.

5.1 Design background

The main flow of summarization and visualization for trend information is obtaining the user's query, collecting and extracting the related data or materials, which are portions of articles in the case of textual data, and then analyzing the data/materials for summarization. Thus, for our studies we need examples of users' queries and examples of the summarizations that the system generates. In addition, the extracted portions of text accompanied by a proper level of analysis are important intermediate results. Such data can behave as the hub for several components of summarization. In order to design such a data set, we must consider more concretely the kind of queries that would be input and the kind of data or materials that would be needed to answer such queries.

Queries on trends input by users can be divided into two groups. This classification is related to how the queries are "clipped."

Clipping by statistics or event types as a whole A user may input a query about the trend in the price of gasoline, which could be explained by a report on two statistics: the nationwide average of the pump price of gasoline and the Dubai oil price. Another user may input a query about the landfall of typhoons, which could be explained by a report compiling dates and places of landfall over the past year. These queries, for prices and landfall events during a specified period, are considered as requests for some statistics or some event types as a whole.

Clipping by agent Users may input a query about the trend in Sony's business or the trend in the Koizumi cabinet. In these cases, their requests do not concern specific statistics but instead involve several statistics about a specific agent. Precisely speaking, the trend in Sony's business is the trend in corporate performance at Sony, which can be shown by several related statistics, such as net sales, operating profit, ordinary profit and so on. So in this case, the user's query covers a set of statistics and is narrowed down to those portions related to a specific agent, which is Sony in this case. For another example, when a user inputs a query on home run tallies in major league baseball for 1998, the query is narrowed down to the home run events of Sammy Sosa and Mark McGwire.

Text information used in order to answer these queries, that is, the materials for summarization and visualization, can be categorized into the following two types.

Description of a statistical value at a time For the cases of trends in gasoline prices and the corporate performance of Sony, each piece of information consists of a statistical value at a certain time point or period. A comparison to previous values is often included, such as "The price of gasoline (one liter, regular) reached a national average of 92 yen; 1 yen higher than the average price of the last week," or "According to the March 1998 financial statements announced on April 7, Sony's combined consolidated profit reached 453 billion 739 million yen (up 45.2% from the previous year)." In this case, summarizing the changes in statistical values could be achieved directly from those pieces of information.

Description of a specific event For the cases of trends on typhoons and home run tallies, the majority of information consists not of descriptions of statistics but rather descriptions of a specific event. Typical materials are "Medium-strength typhoon No. 10 struck Makuraszaki-shi, Kagoshima, at about 4:30 pm on the 17th, and will strike in the vicinity of Shukumo-shi, Kochi the same night." or "On the 25th, Sosa hit his 66th home run, and leads McGwire in the home run tally." Statistical descriptions such as the total number of typhoons that made landfall that year are rarely seen. There are some cases, however, in which the description of an event accompanies a comment on statistical values, as in "Sosa hit his 54th home run; only 7 more to reach Maris record of 61 home runs." In this case, in order to compile a trend, an extra process is needed to obtain a statistical description from the event descriptions. For example, the number of typhoon landfalls in Shikoku and the number of home runs in September should be calculated from those descriptions.

The data set should cover the two-way combination when selecting examples of users' queries.

Gasoline price (20) [statistics as a whole]
Nationwide average of pump price of gasoline [statistic]
Dubai oil price [statistic]

SONY (9) [narrowed down by agent]
Consolidated net sales [statistic]
Consolidated operating profit [statistic]
Consolidated ordinary profit [statistic]
Consolidated net profit [statistic]

Personal computer industry (20) [statistics as a whole]
Domestic shipment volume [statistic]
Domestic shipment value [statistic]
Market share of shipment volume of major makers [statistic]

Unemployment rate (14) [statistics as a whole]
Number of unemployed people [statistic]
Number of unemployed people classified by cause [statistic]
Total unemployment rate [statistic]
Total unemployment rate by gender [statistic]

Typhoon (15) [event types as a whole]
Landfall (place and date) [event]

Home run tallies of major league baseball (13) [narrowed down by agent]
Hitting a home run (by Sosa or McGwire) [event]
Number of home runs in the season (by Sosa or McGwire) [statistic]

Figure 2. Examples of Toics, Statsitics, and Events

5.2 Specifications

Based on the considerations above, a data set was constructed and provided to the participants. The document set consists of Japanese Mainichi newspapers from 1998 and 1999. We selected 20 topics as examples of users' queries clipped by statistics or events or agents, and selected related statistics and event types, around 3 for each topic, 58 in total. These correspond to the input examples and the results of the first step of the framework shown in Figure 1. Examples of topics and selected statistics and event types are shown in Table 2, which also shows their categorization. Numbers in parentheses indicate the number of articles collected.

For each topic, we collected and constructed the following data:

1. List of newspaper articles concerning the topic
2. Annotations on those articles
3. Examples of textual summaries on the selected statistics and events
4. Tables or charts on those statistics

List of articles provides proper examples of articles obtained through the step for information collection in the architecture shown in Figure 1. *Textual summaries* is the output of the step for summarization of each statistic, *Tables and charts* plays two roles: data collected for summarization and/or visualization, and examples of summarization results. Note that these summaries and charts may not have appropriate relevance since they do not take users' queries into account. They are used only as examples for reference, not as the model to be followed.

Annotated articles is the most important data. It represents the intermediate results of the summarization step; that is, the annotations correspond to the results of extracting important sentences and the semantic analysis and context processing of those sentences. The annotation was designed so that its result provides useful information for following the summarization step, and marks the names and the values of statistics, dates, parameters, and so on. Values are annotated so that it is clear for which statistics it is. Dates are annotated with absolute dates when using relative expressions. For example, "Yesterday" is annotated with "19990203." Examples in which the annotation schema was applied to the English texts are shown in Figure 3. Examples taken from the data set are shown in Figure 4. Specifications for annotation are explained in the Appendix. In the flow of summarization in general, extracting important sentences and analyzing them is followed by rephrasing and sentence generation to eliminate redundancy and maintain consistency. This annotation corresponds to the output of the former and the input of the latter. For researchers interested in important sentence extraction or language processing for summarization or information extraction, it can be referred to as the correct result of their process, and for researchers interested in the latter process such as rephrasing and sentence generation, it becomes input data, in which several fundamental analyses,such as named entity extraction and temporal processing [5] are already completed. In this sense, the annotated articles behave as a hub of the summarization step.

6 Course, Current Situation, and Future Plans

MuST was announced in November 2004, at the joint meeting of SIGNL of the Information Processing Society of Japan and SIGNLC of the Institute of Electronics, Information and Communication Engineers [3]. Mailing lists on related topics were used to explain the keynote and to collect researchers interested

<unit stat="nationwide average of pump price of gasoline"><name part="head">The price of gasoline (one liter, regular)</name> reached <name part="foot">a national average of</name> <val>92 yen</val>; <rel>1 yen</rel> higher than the <name part="foot">average price</name> of <date gra="week" abs="19990610">last week</date></unit>.

<unit stat="combined consolidated profit">~~According to the March 1998 financial statements announced on April 7,~~ <topic>Sony</topic> 's <name>combined consolidated profit<name> reached <val>453 billion 739 million yen<val> (up <rel>45.2%<rel> from <date gra="year" abs="1997">the previous year</date>)</unit>.

<unit evet="typhoon landfall">Medium-strength <par>typhoon No. 10</par> struck <par> Makuraszaki-shi, Kagoshima</par>, at <date gra="hour" abs="19980917">about 4:30 pm on the 17th</date>, and will strike in <par>the vicinity of Shukumo-shi, Kochi</par> <date gra="hour" abs="19980917">the same night</date></unit>.

<unit stat="domestic shipment volume"><name>The domestic shipment volume</name> for <date gra="half-year" abs="199804">the first half of the year</date> was <val> 4,391,000</val>, which is <rel>34%</rel> higher than in <date gra="half-year" abs="199704">the same period last year</date> and marked <rel>the highest record</rel> for the half-year range</unit>.

Figure 3. Examples of Annotation on English text

990617020 **Gasoline price**
<unit stat="レギュラーガソリン全国平均店頭価格">~~石油情報センターが１６日発表した給油所石油製品市況調査によると、~~<date gra="週" abs="19990617">今週</date>調査の<name part="head">ガソリン価格（レギュラー１リットル）</name>は<name part="foot">全国平均</name>で<val>９２円</val>となり、<date gra="週" abs="19990610">前週</date>の<name part="foot">平均</name>に比べ<rel>１円</rel>上昇した</unit>。<unit stat="レギュラーガソリン全国平均店頭価格"><name part="foot">全国平均</name>の上昇は<date gra="旬" abs="19990201">２月上旬</date>以来、<dur gra="月">４カ月</dur>ぶりだ</unit>。

980122071 **Personal computer industry**
<unit stat="メーカー毎出荷台数">メーカー別では、<dur gra="月" abs="1997">年後半</dur>から<par>ソニー</par>のノートパソコンが急伸</unit>。<unit stat="メーカー毎出荷台数シェア"><par>ＮＥＣなど昨年の上位５社</par>の<name>シェア</name>は<pro ref="前年度比" id="980122071_1">同</pro><rel>３・１ポイント</rel>低い<val>８２・７％</val>となった</unit>。<unit stat="メーカー毎出荷台数シェア"><par>アップルコンピュータ</par><ins>の<name>シェア</name></ins>は経営危機のニュースなどで<dur gra="月" abs="1997">年前半</dur>に苦戦し、<pro ref="前年度比" id="980122071_1">同</pro><rel>２・８ポイント</rel>下落して<val>７・８％</val>と<val>１０％の大台</val>を割り込んだ</unit>。

981018004 **Typhoon**
<unit event="台風上陸">中型で並の強さの<par>台風１０号</par>は、<date gra="時" abs="19980917">１７日午後４時半ごろ</date><par>鹿児島県枕崎市</par>に上陸、<date gra="時" abs="19980917">同日夜</date>には<par>高知県宿毛市付近</par>に再上陸して進み~~、西日本を縦断した~~</unit>。

Figure 4. Examples of Annotated Articles

in its subjects. At the beginning of 2005, we opened a mailing list for interested persons and also started to call for participants. At the same time, the data set, excluding tables and charts on statistics, was delivered to the participants[1]. The tables and charts on statistics were delivered in April in the form of a URL list linked to charts and tables on the Internet. After that, there were two major revisions on the annotation of articles: a specification change in May, and extensive revision and correction of errors in August. The participants were also able to use the Mainichi newspaper corpus since MuST is a pilot workshop of NTCIR. In August, a new mailing list, closed to the current users of the data set, was started in addition to the previous

[1] www.kecl.ntt.co.jp/scl/workshop/must/must_index.html

one for interested people, in order to encourage more substantial discussion of research on the data set. This new ML is used mainly for discussing questions on the data set, while the other is used for announcements from the organizers.

MuST had thirteen participants as of October 2005. The number of participants on the ML for interested people is over 60.

We will hold a closed workshop in March 2006, where all participants will give a presentation on their research progress. As opportunities for research presentations to the public, a session on related themes will be set up at the NLC Symposium in March 2006, and a workshop on related themes is proposed for the annual meeting of the Association for Natural Language Processing in March 2006. In addition, a special issue on "visualization and summarization of text" will be published by the Japan Society for Fuzzy Theory and Intelligent Informatics in October 2006; the deadline for submission of papers is the end of February 2006. We are also planning a round table meeting in November 2005 to encourage discussion among the participants.

7 Conclusion

This paper reported on the MuST (Multimodal Summarization for Trend Information) workshop, which was designed to encourage cooperative and competitive studies on summarization and visualization for trend information. A noteworthy feature of the workshop is that the participants share the same research resource, which constitutes the centripetal force of the workshop. MuST started at the beginning of 2005, and the first period will come to an end in March 2006. Looking toward the next period, efforts are being made to establish the course for encouraging further research in this field.

Acknowlegements

The workshop and related research are supported in part by grants from a joint research of Nippon Telegraph and Telephone Corp. and The University of Tokyo, and a joint research of National Institute of Informatics and The University of Tokyo.

References

[1] Jun'ichi Fukumoto, Tsuneaki Kato and Fumito Masui. Question Answering Challenge (QAC-1) An Evaluation of question answering tasks at the NTCIR workshop 3. *AAAI 2003 Spring Symposium New Directions in Question Answering*, pp. 122-133, 2003.

[2] Tsutomu Hirao, Takahiro Fukusima, et al. Corpus and Evaluation Measures for Multiple Document Summarization with Multiple Sources. *Proceedings of 20th International Conference on Computational Linguistics*, pp. 535 – 541, 2004.

[3] Tsuneaki Kato, Mitsunori Matsushita, and Tsutomu Hirao. Proposal of a Workshop on Summarization and Visualization of Trend Information (in Japanese). *IPSJ SIGNL 2004-NL-164*, pp. 89–94, 2004.

[4] Tsuneaki Kato, Mitusnori Matsushita, and Noriko Kando. Summarization and Visualization on Trend Information – Research Issues and Workshop – (in Japanese). *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics* , Vol.17, No.4, pp.424-431, 2005.

[5] Inderjeet Mani and George Wilson. Robust Temporal Processing of News. *Proceedings of ACL-2000*, pp. 69–76, 2000.

[6] Manabu Okumura. Automatic Text Summarization (in Japanese). *IPSJ Magazine*, Vol. 45, No. 6, pp. 574-579, 2004.

[7] Satoshi Sekine. Information Extraction; Extract Structural Information from Unstructured Text (in Japanese). *IPSJ Magazine*, Vol. 45, No. 6, pp. 569-573, 2004.

[8] Russell Swan and James Allan. Automatic Generation of Overview Timelines. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 49–56, 2000.

[9] Koichi Takeda and Hiroshi Nomiyama. Information Retrieval using Textual Information Visualization (in Japanese). *IPSJ Magazine*, Vol. 41, No. 4, pp 343-350, 2000.

[10] Ellen M. Voorhees and Dawn M. Tice. Building a Question Answering Test Collection. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 200 – 207, 2000.

[11] Michael White, Tanya Korelsky, et al. Multidocument Summarization via Information Extraction. *Proceedings of Human Language Technology 2001*, 2001.

APPENDIX

Specifications for Annotation

`unit` **element** Portion, sentence in principle, referring to a specific statistic or event. What statistic or event that portion is referencing is shown

using the following attribute: `stat` for statistics and `event` for events. When the portion states a prospect rather than a fact, it is shown by `type = "pros"`. All other annotations are made inside this element with the exception of the `rft` element.

`del` **element** Portion, clause in principle, which does not include information on the value or change in the statistic or event concerned although it is in a `unit` element. Sometimes states a news source or basis, or subjects related to the undeleted part of the `unit` element.

`ins` **element** Portion supplied in order to make a `unit` element understandable on its own. It is mainly the name of a statistic or a parameter followed by a proper postposition that was omitted in the original text because it is recoverable from the content outside of the `unit` element. Insertion of this element must be made in such a way that the resultant expressions are well-formed Japanese sentences.

`name` **element** Name of statistic. The `stat` attribute is used for specifying the statistic it stands for, and can be omitted when the `unit` element containing this element concerns only one statistic. In some cases, the name of the statistic is distributed throughout a sentence, such as in "the gasoline price is 101 yen for the nationwide average" for "the nationwide average price of gasoline." These parts are annotated with `part = "head"` for the semantically main part and `part = "foot"` for the auxiliary or modified part.

`val` **element** Value of statistic. This element also has the `stat` attribute as in the case of the `name` element.

`rel` **element** The difference or ratio or order of the value of a statistic, which is related to the value of a statistic but is not the value itself. This element also has the `stat` attribute as in the case of the `name` element.

`date` **element** Description of date and time such as "10th," "this year," and "yesterday." This element has the `gra` attribute, which indicates the granularity or particularity of that date, such as day, month or year, and the `abs` attribute, which indicates the absolute time that the expression denotes. The absolute time is expressed with four digits when it has year granularity and eight digits when it has day granularity.

`dur` **element** Description of time period and interval, such as "five years" in "the average price reached 100 yen after an interval of five years." Parameters of statistics, such as "this month" in "the average price of this month," are not this element, but the `date` element.

`par` **element** Parameter of statistic or event except the date, such as the subject of a statistic or the agent of an event or the place of occurrence. Company names for market share and locations of typhoon landfall are examples of this element. Agents specified as users queries, such as Sony and Sosa, are specially treated as `topic` elements.

In addition to the above, the annotation has a `cond` element, which states some supplementary conditions on the described statistics and events, and `pro` and `rft` elements, which involve reference expressions.