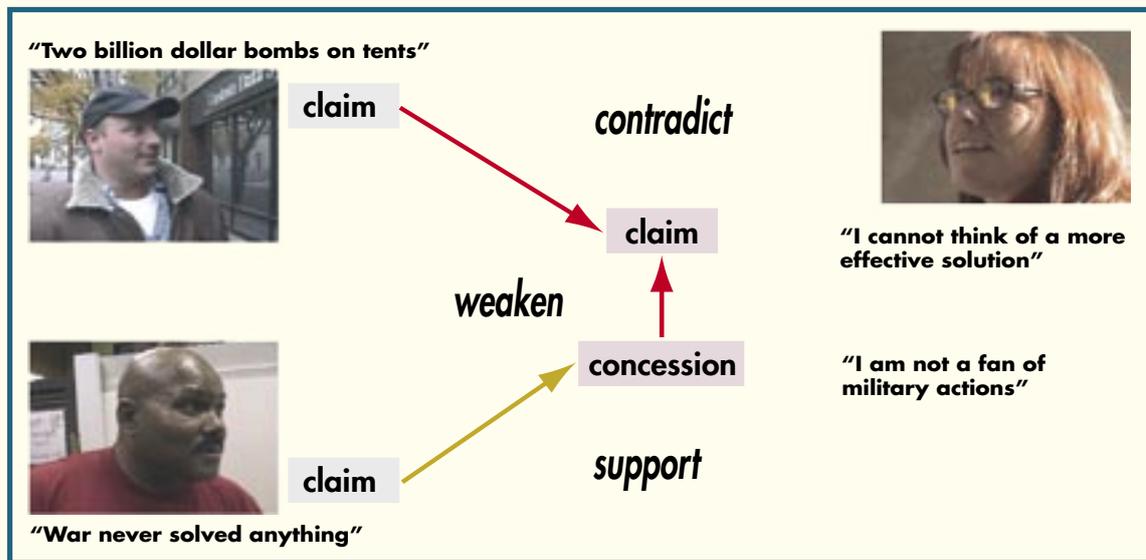# VOX POPULI SYSTEM AUTOMATICALLY GENERATES DOCUMENTARIES

By selecting and assembling video fragments, a documentary maker by definition projects her personal vision and preferences on the material at hand. Vox Populi, a *generation engine* developed by the research group Multimedia and Human-Computer Interaction, replaces this subjective work of the documentary maker by an automated process.

The amateur filmmakers of *Interview with America,* a collection of interviews with Americans from different socio-economic groups about the immediate consequences of the terrorist attacks of September 11, shot eight hours of video material. The makers could have put all this material onto the internet in the form of a database. The downside of such an approach is evident: the extent of the material is such that few would take the time to go into it. An alternative would

sidestep these disadvantages. The Vox Populi system draws from an audio-visual database – for example, *Interview with America* – to automatically generate short video documentaries. The intention is that these generated video sequences follow a logical line of argumentation and draw a conclusion. With this procedure all stored material remains easily accessible without the direct interference of an editor, and nevertheless the presentation forms a developing unity guided by argumentation structures.

To automatically generate video fragments under these circumstances, several steps have to be taken. First of all a method is needed that recognizes that there are various approaches to a subject and that a variety of conclusions can be drawn. For this purpose Vox Populi uses the so-called Toulmin model, which dissects the structure of a rational discourse into specific parts: the claim, the data, the backing, the concession and the

"Two billion dollar bombs on tents"

claim

contradict

claim

"I cannot think of a more effective solution"

weaken

concession

"I am not a fan of military actions"

claim

support

"War never solved anything"

An example from the 'Interview with America' database (IWA).

be to turn the material into a short documentary. This approach, however, has the disadvantage that much potentially interesting material will get lost and that the presentation is coloured by personal preferences and opinions.

With its prototype *generation engine* Vox Populi, the research group Multimedia and Human-Computer Interaction (INS2) has developed an instrument to

condition. The claim is supported by the data and the backing, and it is weakened, differentiated or argued by the concession and the condition.

The documentary maker identifies statements in the database in the form of video segments, each given a specific annotation. An annotation is a formal description of what is on the video (encoded in this case in a standard web format RDF), using terms

from a limited vocabulary the computer understands. Two types of annotations are used: descriptive and rhetorical. Descriptive annotations answer the who, what, where, when questions in that particular segment. They describe, for example, who made the statement and this person's background. This enables the user to give Vox Populi the following assignment: 'Select all answers to question X, given by people of race Y and education Z.'
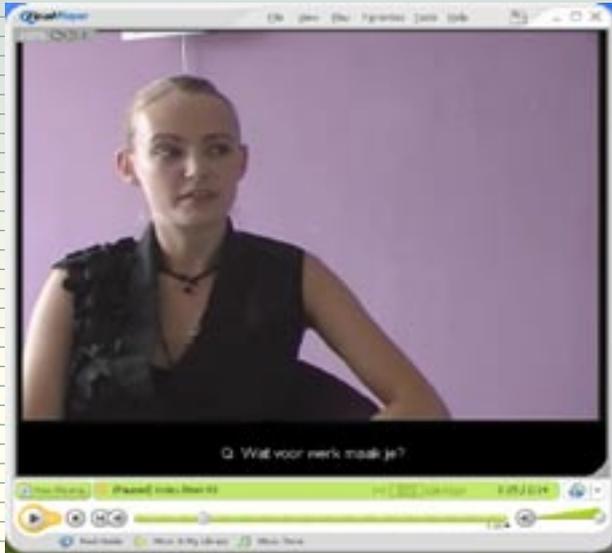
The Vox Populi software is used by the Netherlands for Media Art Institute Montevideo/Time Bases Arts to reproduce video interviews with VJs.

Rhetorical annotations, as a next step, enable the system to present these answers in an argumentation structure. An example of a statement, from *Interview with America*, is the following: 'I am not a fan of military actions, in the big picture I don't think they are ever a good thing.' A statement or video segment – with a self-reliant meaning– is first encoded using a three part syntactical structure. In the example: *military actions* (subject) *never* (modifier) *positive* (predicate). The terminology is part of a domain dependent vocabulary. The documentary maker can develop it personally or she can use an online lexical reference system such as Wordnet.

The system can manipulate the statements using the relations between the terms. The documentary maker can choose between four relations: *similar, opposite, generalization and specification*. Thus 'war' is similar to 'military actions' and is opposite to 'diplomacy'. These relations complete the rhetorical annotations.

With such a fully completed annotation scheme Vox Populi can generate alternative statements, based on the relations between used terms. Thus the system recognizes the statement 'War is the best solution' as support of the statement 'military actions are the best solution', because the terms 'military actions' and

'war' are annotated as *similar*. For the same reason the system concludes that the statement 'diplomacy is the best solution' *contradicts* the earlier statement. In principle, an even number of opposite relations should produce a similar relation, therefore a support of the initial statement.

To determine the final video sequence of the 'documentary', Vox Populi offers a choice between various strategies. The system can choose a fragment of an interview (a 'segment') wording a statement that is subsequently contested by an alternative fragment. Such segments can support the concession of the interviewee (through the relation *similar*) or contradict the claim (through the relation *opposite*). The result is a rebuttal of the original statement. This way a statement can just as easily be supported or differentiated, based on the available segments.

As the documentary maker cannot foresee all possible generated video sequences, her personal vision is detached from the reasoning in the final presentation. This can be a revolutionary change in the way documentary makers are used to work. For a broad acceptation of Vox Populi, however, it is essential to develop tools that integrate the labour-intensive annotation into the production environment.

Vox Populi was developed by CWI's INS2 research group, which is specialised in methods for automatically generating interactive multimedia documents. In addition to Vox Populi, the group has developed other presentation-generating systems, such as 'Topia', based on an annotated database of art works from the Rijksmuseum in Amsterdam. Vox Populi is financed by the NWO project ToKeN/I²RP. The group is also involved in Bsik MultimediaN/E-Culture, IST NoE K-Space, ITEA/Senter Passepartout, NWO ToKeN/ CHIME and NWO NASH.

For more information:
www.cwi.nl/ins2

Prof.dr. L. Hardman
Telephone: +31 20 - 592 4147
E-mail: Lynda.Hardman@cwi.nl

November 2005