K

KNOWLEDGE-BASED MULTIMEDIA ADAPTATION DECISION-TAKING

Dietmar Jannach[a], Christian Timmerer[b], and Hermann Hellwagner[b]
[a]Department of Computer Science, Dortmund University of Technology, Germany
[b]Department of Information Technology, Klagenfurt University, Austria

Synonym: *Artificial Intelligence (AI)-based techniques for finding a sequence of adaptation operations; searching for a sequence of adaptation operations utilizing AI-based planning techniques.*

Definition: *Knowledge-based multimedia adaptation decision-taking is referred to as the process of automated construction of a suitable sequence of adaptation operations for a given multimedia resource and a set of environmental requirements based on semantic annotations.*

Problem Description

In the context of personal delivery of multimedia content over the Internet one of the main challenges lies in the fact that the provision of an end-user multimedia service has to be done in a heterogeneous and constantly evolving environment. End users are using different types of client devices with different capabilities, for instance with respect to display sizes and the network types they support. In addition, recent standards in the field like MPEG-21 [1] also provide for mechanisms that allow end users to specify explicit, personal preferences, which should be taken into account when servicing the client. On the other hand, not only the number of file and encoding formats, in which multimedia resources are stored at the server constantly increases, but also the new possibilities of annotating the resources, for instance based on MPEG-7 [2], open new opportunities for personalization on the content level.

In order to take these preferences and personalization requirements into account, a multimedia server will therefore adapt the multimedia resource correspondingly before sending it to the client, in particular because the capabilities of the client's device of transforming the resource by itself may be limited. Given the large number of possible media formats and the new opportunities of personalizing the content and the consumption style, the problem arises that the adaptation on the server side cannot be done in a single step or with the help of a single media transformation software tool. An intelligent multimedia server will thus have to apply a *sequence of transformations* on the original media and also use tools or libraries from several, potentially highly-specialized software providers.

This process of determining a suitable set of operations that transform a given media resource according to the requirements of the user and other environmental parameters like the network bandwidth is referred to as adaptation decision-taking.

The question now arises how to implement the logic of a corresponding Adaptation Decision-Taking Engine (ADTE) such that it

a) is capable of computing transformation sequences of arbitrary length, given some user requirements, environmental parameters, and a description of the existing media resource,
b) guarantees that the chosen sequence will result in a correctly adapted resource,
c) is easily extensible when new software tools, user preference types, or media formats become available.

Finally, another desirable property of such a mechanism is that it is compatible with and conforms to existing standards such that transformation tools of different vendors can be easily integrated.

Figure 1 summarizes the problem setting and in particular shows where the *adaptation decision-taking engine* is placed in a general architecture of an intelligent multimedia server. The inputs to the ADTE are i) the *usage environment description* containing the user requirements and other environmental parameters, ii) *content descriptions* that carry information about the original resource, and iii) *tool descriptions* that describe the available adaptation operations. The result of the adaptation decision-taking process is called the *adaptation plan*, which consists of a list of transformation operations and which can also include a set of parameters for each operation. The *adaptation engine* actually invokes the individual tools mentioned in the plan which transform the original media according to the usage environment.

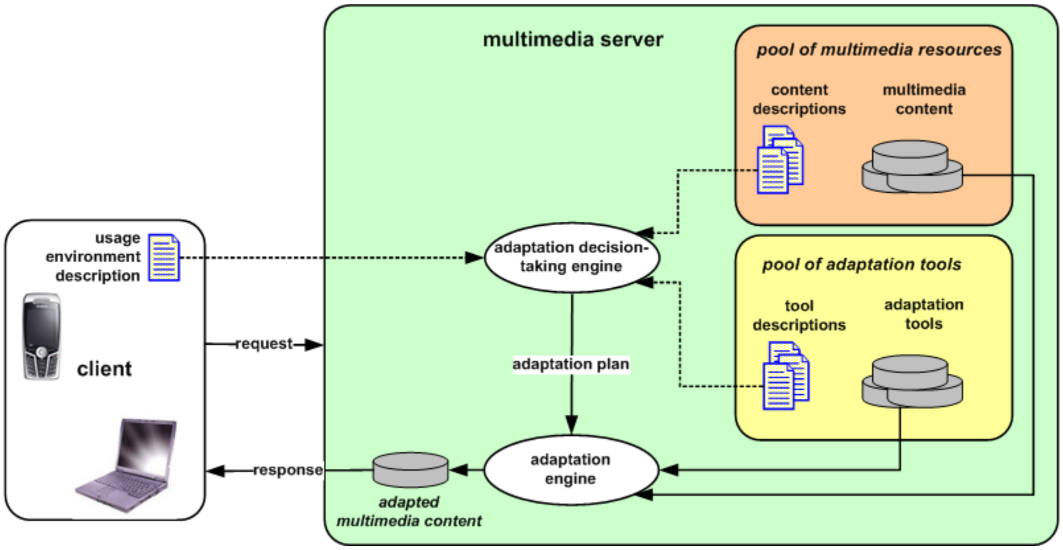

Figure 1. Adaptation Decision-Taking Architecture (adopted from [3]).

Technical solution approaches

In order to address the above-mentioned requirements, *knowledge-based approaches* to adaptation decision-taking have been proposed for instance in [3], [4], and [5]. The main idea of these proposals is to exploit the well-known advantages of the knowledge-based system development paradigm. The advantages include the declarative form of knowledge representation, which allows for easy maintenance and extensibility as well as the opportunity to exploit or re-use problem solving algorithms which are independent of the domain and the particular scenario. Additionally, the approaches described in [3-5] rely on recent XML-based standard formats for knowledge representation and in particular exploit new mechanisms for semantically annotating the functionality and effects of applying a specific transformation tool.

Subsequently, we will discuss details of the KoMMa system described in [3] and [4] as this approach also addresses the problem of the technical integration of the chosen knowledge representation scheme with the existing ISO/IEC MPEG standards (see [6] for details of an ISO/IEC Core Experiment in that area).

Within the KoMMa system, the main idea is to view the problem of generating the required sequence of adaptation operations that transforms an existing multimedia resource according to the given requirements as a *state-space planning problem* (see [7] for an introduction). State-space planning is a classical technique from the field of Artificial Intelligence (AI) and can be summarized as follows. The inputs to the planning problem are a description of the current state of the world, a description of the desired world state, as well as a set of *world-altering* actions which can be executed by an intelligent agent. The planning problem consists of determining a sequence of parameterized actions that transform the world from the existing state to the target state, a problem that has been extensively studied over the last decades and for which various highperformance, domain-independent algorithms are available today. For describing the initial state and the target state, propositional logic can be used; the actions are described in terms of their inputs, outputs, preconditions, and effects (IOPE).

The similarity of the adaptation decision-taking problem which we described earlier and the state-space planning problem can be easily seen. That is, the goal is to transform an existing media resource (i.e., the start state) into the format desired by the client (i.e., the goal state) by use of a set of transformation operations (i.e., the set of actions). An example of such a simple planning problem that includes the operations "grey-scaling", "spatial-scaling" as well as "encoding/decoding" for adapting a color video for a black/white mobile device with a small display could be the following.

Start state: description of existing resource	*Goal: user preferences, device capabilities*
coding_format(mpeg_4).	decoding_capabilities(mpeg_1).
color_domain(true).	color_capabilities(false).
frame_size(640,480).	display_size(320,240).

The spatial-scaling and grey-scaling actions could be described as follows.

Action: spatial-scaling	
Input:	image, x, y, newx, newy
Output:	spatialScaledImage
Preconditions:	yuvImage(image), width(x), height(y)
Effects:	yuvImage(spatialScaledImage), width(newx), height(newy), horizontal(newx), vertical(newy)

Action: grey-scaling	
Input:	image
Output:	greyScaledImage
Preconditions:	yuvImage(image), color(true)
Effects:	yuvImage(greyScaledImage), color(false)

An "adaptation plan" computed by any state-space planner could be the following. Note that the symbols *fb1* to *fb5* are used to forward a symbolic pointer to the media stream from one operation to the next.

1:	decode(fb1, mpeg_4, fb2)
2:	spatial-scaling(fb2, 640, 480, 320, 240, fb3)
3:	grey-scaling(fb3, fb4)
4:	encode(fb4, mpeg_1, fb5)

Note that the above listings are merely used to sketch the basic approach in a proprietary notation. If we however want to ensure interoperability and extensibility of the approach, a few more aspects have to be considered. First, the question arises how we can make sure that a *consistent set of (predicate) symbols* is used to describe the start and the goal state as well as the actions, which would allow us to easily integrate third-party tools. In addition, we also have to consider aspects of the *representation format* and finally, how the described knowledge can be smoothly integrated with the data representation formats of existing standards such as the XML-based MPEG-21 standard.

These aspects are covered in the approach described in [3] and [4] as follows. The problem of using a (strictly) defined vocabulary is resolved by interpreting the existing MPEG-7 and MPEG-21 standards as the *domain ontology*. MPEG-7 for instance specifies how the given resolution or the color domain of a multimedia resource has to be described. With the help of MPEG-21 Digital Item Adaptation [9] usage environment descriptions one can, on the other hand, for example describe the capabilities of the client device. Thus, as only defined terms are used, the compatibility of the descriptions in the knowledge base of actions and states can be guaranteed. Moreover, when looking at the listing describing the *spatial-scale* above, one can also see that this form of representation also allows for an easy mapping of MPEG-7 related aspects like *width* and *height* of a frame to MPEG-21 elements like the *horizontal* and *vertical* resolution of the end user device.

With respect to the representation format, [3] and [4] rely on OWL-S[1], a language (more precisely an ontology) for semantically annotating Web Services. Similar to classical AI planning approaches, OWL-S follows an IOPE-style modeling approach, such that standards planning algorithms can be exploited to automatically construct complex composite services from atomic Web Services. With regard to data representation, OWL-S also proposes an XML-based syntax, which on the one hand allows for the integration with the underlying Web Service technology (like WSDL[2]) and on the other hand provides mechanisms to add references to ontological definitions in other XML documents. The full technical integration starting from the MPEG standards over OWL-S process definition down to the concrete binding in WSDL and Java types as well as a reference implementation are described in detail in [4].

Comparison and Open Issues

The work described in [5] is to some extent similar to the work described in [3] and [4] as it also relies on the usage of a defined domain ontology and a planning algorithm for the task of Web Service composition. Compared to the approach described above, [5] however uses an additional, manually defined ontology which is modeled after the elements of different standards. Additionally, [5] also uses a specific type of planning algorithm based on *hierarchical task network planning*. In general, it can however be seen in both approaches that knowledge-based approaches can help to build more flexible and "intelligent" multimedia adaptation servers as it would be possible with previous approaches, which are for instance based on manually engineered *adaptation chains* like described in [8].

With respect to open research issues, the following aspects which are relevant in practical scenarios have not been addressed in the above-mentioned approaches yet. First, the planning algorithms of [3-5] will – due to their property of soundness and completeness – guarantee that if a transformation sequence can be found, it will properly transform the

---

[1] http://www.w3.org/Submission/OWL-S/
[2] http://www.w3.org/TR/wsdl

media from the original format to the desired one. Still, there is no guarantee that the identified transformation plan is a good or optimal one with respect to some cost model. For instance, if we think of a plan that consists of shrinking the image and removing the color, it could be better to first reduce the image size before applying any other operation on it. The practical experiments discussed in [4] show that typical plans are rather short and that further optimization would most probably not be a problem with respect to computational complexity. The question for future work, however, is how *plan costs* or other heuristics can be modeled in an interoperable and comprehensible manner in such a knowledge-based framework.

Another possible area of future work is the question of how to cope with situations in which the decision-taking engine fails to determine a plan which fulfills all the constraints and preferences of the client. One obvious goal in such situations could be to search for plans that result in an adapted media that fulfills the client's constraints *as good as possible*. For accomplishing such a task, the used planning technology has to be extended in a way that it can also cope with what is commonly called *soft constraints*. Additionally, also an interoperable form of specifying soft and hard constraints as well as client-specific priorities has to be developed.

Finally, it is also open to some extent how such a proposed knowledge-based multimedia adaptation framework shall interact or can be integrated with the other adaptation-related layers and technologies of MPEG-21 such as Adaptation QoS or – at the lower level of media adaptation – Bitstream Syntax Descriptions [9].

See: Adaptation Decision-Taking, Optimization-based Multimedia Adaptation Decision-Taking, MPEG-21 Digital Item Adaptation, MPEG-21 Multimedia Framework

References

1. I. Burnett, R. V. de Walle, K. Hill, J. Bormans, and F. Pereira, "MPEG-21: Goals and achievements", *IEEE MultiMedia Magazine*, vol. 10, no. 6, pp. 60–70, March 2003.
2. J. M. Martinez, R. Koenen, and F. Pereira, "MPEG-7—The generic multimedia content description standard Part 1", *IEEE MultiMedia Magazine*, vol. 9, no. 2, pp. 78–87, April-June 2002.
3. D. Jannach, K. Leopold, C. Timmerer, and H. Hellwagner, "A Knowledge-based Framework for Multimedia Adaptation", *Applied Intelligence*, vol. 24, no. 2, pp. 109-125, April 2006.
4. D. Jannach and K. Leopold, "Knowledge-based Multimedia Adaptation for Ubiquitous Multimedia Consumption", *Journal of Network and Computer Applications*, vol. 30, no. 3, pp. 958-982, August 2007,.
5. M. De Geyter and P. Soetens, "A Planning Approach to Media Adaptation within the Semantic Web", *Proceedings of the 11th Intl. Conference on Distributed Multimedia Systems (DMS 2005)*, Banff, Canada, pp. 129-134, September 2005.
6. C. Timmerer, K. Leopold, D. Jannach, H. Hellwagner, and M. Kimiaei, "Report of CE on Conversion Parameters", *ISO/IEC JTC 1/SC 29/WG 11 M11262*, Palma de Mallorca, Spain, October 2004.
7. S. Russel and P. Norvig, *Artificial Intelligence: A modern approach*, Prentice Hall, 2003.

8. L. Böszörmenyi, H. Hellwagner, and P. Schojer, "Metadata-driven Optimal Transcoding in a Multimedia Proxy", *Multimedia Systems*, vol. 13, no. 1, pp. 51-68, September 2007.
9. A. Vetro and C. Timmerer, "Digital Item Adaptation: Overview of Standardization and Research Activities", *IEEE Transactions on Multimedia*, vol. 7, no. 2, pp. 418–26, June 2005.