TransRouter revisited - Decision support in the routing of translation projects

Rainer Hammwöhner
Department of Information Science
University of Regensburg
93040 Regensburg
rainer.hammwoehner@sprachlit.uni-regensburg.de

Abstract

This paper[1] gives an outline of the final results of the TransRouter[2] project. In the scope of this project a decision support System for translation managers has been developed, which will support the selection of appropriate routes for translation projects. In this paper emphasis is put on the decision model, which is based an a stepwise refined assessment of translation routes. The workflow of using this system is considered as well.

1. Introduction

Quality and efficiency of translation services depends on the appropriate routing of translation projects. The choice of a translation route requires knowledge about the technology at hand, the resources available and the relevant project parameters. For several reasons this information is difficult to obtain for translation managers.

- Translation technology is undergoing a swift development process. As a consequence new approaches to translation support or new systems will not be considered in decision making.

---

[1] A more comprehensive description of the prototype is provided by [Hammwöhner 00].

[2] TransRouter [King 99] is project LE4-8345 in the Telematics Applications of Common Interest programme of the Fourth Framework Programme, supported by the Commission of the European Communities and by the Swiss Federal Office for Education and Science. The members of the consortium are:
Berlitz, Dublin (Charles Hughes, John Micks), CST, Copenhagen: (Bart Jongejan, Nancy Underwood), University of Edinburgh (Jo Calder), TIM/ETI, University of Geneva.(Margaret King, Sandra Manzi), Lernaut and Hauspie, Munich (Johannes Ritzke), LRC, Dublin (Keith Brazil, Conor McDonagh, Reinhard Schäler), University of Regensburg (Rainer Hammwöhner, Jürgen Reischer).

- The need of an organisational memory covering the resources (e.g. translation memories, term banks) at hand is not recognised by a majority of translation agencies.
- Relevant project parameters may not be estimated easily. Is the text repetitive and to which degree? Is special terminology employed? Is the text to complex to use a machine translation system?

The objective of the TransRouter project is the design of a decision support system providing the translation manager with the relevant information. Since no widely accepted notion of decision support in this application field can be built on, the general approach of TransRouter is based on the development of a sequence of prototypes that were presented to the public. These prototypes share the following features:

- Several profiles contains the relevant features of agents (translators or translation tools) and resources (translation memories or term banks)

- The features of translation projects are covered by a different type of profiles.

- A set of tools is developed, which allow the (semi)-automatic acquisition of project data (e.g. text size, terminology, complexity, repetitiveness).

- A decision kernel computes translation routes based on agent, resource and project profiles.

The general approach of the TransRouter project has — in an early stage of the project — been presented at the least ISI conference [Hammwöhner 98]. This paper will give an overview on the TransRouter prototype as developed at the University of Regensburg and thus give an account of the final results of the project. The task of route construction and selection consists of a couple of steps, which need specific support each. Thus, a first part of this paper will give an overview on the workflow which may be supported by the system and which is needed to operate it.

Before a detailed discussion can be started, the nature of the intended support should be pointed out[3]. The goal of TransRouter is not to find *the* optimal route for a translation automatically. This approach would need formalised rules of decision making in translation projects, which are not at hand. Furthermore it seems to be questionable whether translation managers would accept a system that would seem to take over the responsibility of decision making. Thus, TransRouter will support the manager in decision making by pointing out alternatives in agent and resource selection and the resulting route choice. TransRouter will give support in the acquisition of the relevant project data, the resource assessment and the assessment of routes with respect to time, costs and quality.

---

[3] An overview on approaches to decision support may be found in [Turban 95]
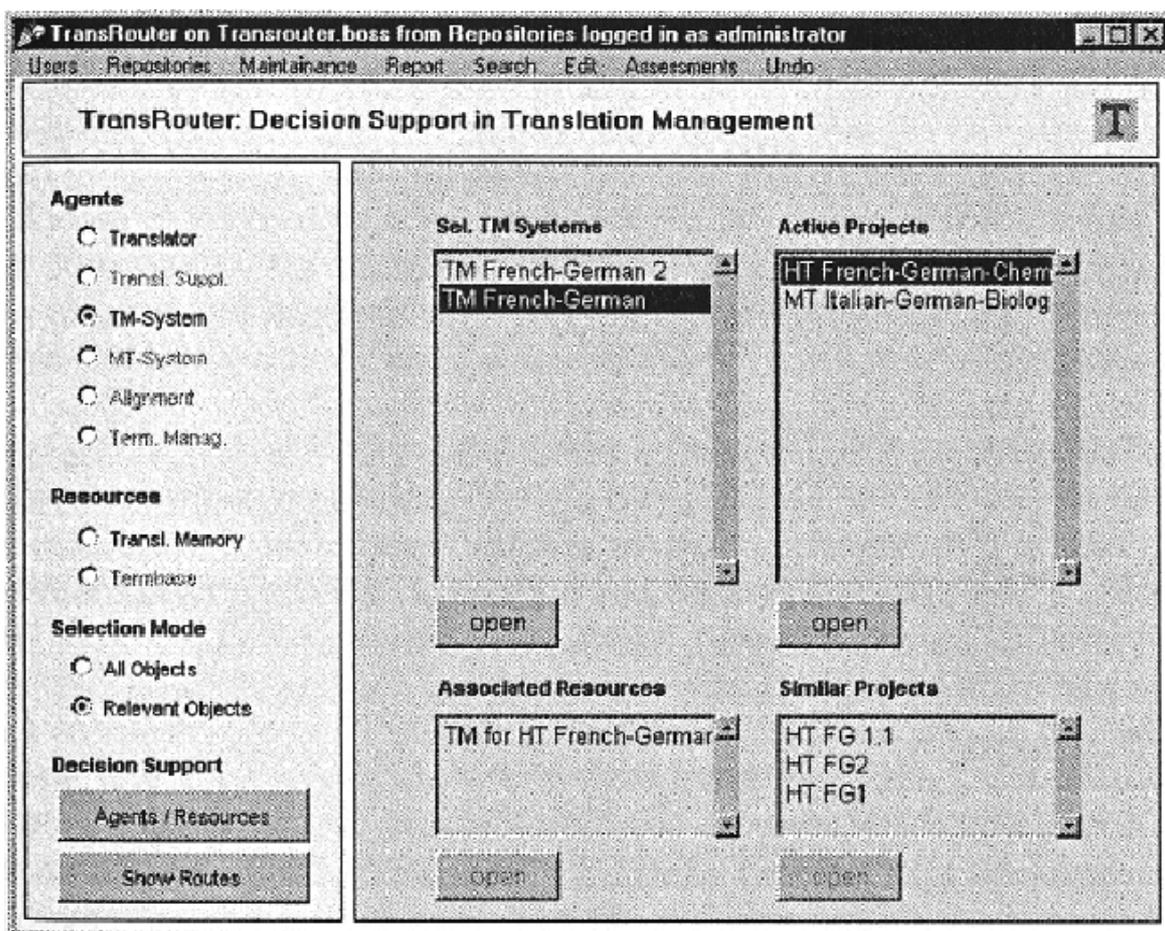
Figure 1: Initial page of TransRouter— lists of projects, agents and resources are provided

2. Workflow

This section will provide a short overview of the general workflow in Trans-Router. A discussion of more detailed workflow will be given then in the following sections.

- Data acquisition: The translation manager will expect agent and re-source data to be available in the system when he starts to use it. Nevertheless there will be a need to update the system at regular intervals.

- Furthermore there should be some means to enter project data conveniently. TransRouter offers appropriate tools for these tasks, but this pa-per will not go into detail here.

- Information retrieval: The user might want to extract data from the system without using the inference mechanism.

- Agent and resource selection: The first step of decision making is the selection of agents and resources, which may be used to process a given

project. This step may be performed automatically since it is based on formally defined criteria. Some manual editing may follow the automatic selection step.

- Route construction: The selected agents and resources are used to construct possible routes, which are based on a built-in route model. The system will not generate all possible routes but a representative set of routes covering all useful route types. The user can refine these routes afterwards.

- Route assessment: The routes, which are created by the system, have to be assessed by the user. He will choose one of several evaluation functions, a cost function, time or quality function or a combination of those. The system will then sort the routes according to those criteria. The manager will then pick one or more routes which seem to be promising. He then may modify the set of tools and resources assigned to these routes.

- Route selection and route processing: when the route assessment and refinement is completed a route can be selected for the further processing of the translation project. At this stage the job of the TransRouter system seems to be done. It will be shown that some additional documentation job will be helpful for the appropriate handling of further projects.

The steps as mentioned above give an outline of a macroscopic workflow in the use of TransRouter.

3. Information Retrieval

The whole process of decision support of course may be seen as some kind of information retrieval. If a user wants to access the TransRouter repository directly he may choose from two retrieval options. Matching oriented retrieval is based on the formulation of a query. This query will be processed by the system, which then will come up with a list of relevant objects fitting the re- quest. Browsing uses implicitly or explicitly given references between objects that can be navigated via the TransRouter interface.

The approach to matching oriented retrieval is based on the construction of sample objects. A user, who wants to find a machine translation system by its specific features, will have to construct an agent profile describing such a system. Truncation symbols (`#' and `*') substituting an arbitrary character or sub-string are available. If attributes are numeric some kind of fuzzy match is performed.

Browsing oriented retrieval employs relationships between objects, which are made explicit by the interface. A profile of a translation memory may provide not only its data storage format but also list translation memory systems, which can process this data format.

One of the most promising applications of browsing in TransRouter seems to be the use of organisational experience. The translation manager will check, whether a machine translation system has already been used in successful projects. Did these projects have anything in common with the project currently in progress? Were there any projects at all, which were similar to the current one? What can be learned from their performance?

This leads to the question of which notion of similarity should be employed. Should the similarity of projects be based on the basic project attributes only or should the chosen routes be considered too? Since the similarity measure should be applicable to all projects — even if newly defined— there must be at least one notion of similarity, which does not take routes into account. The similarity measure should allow a ranking of projects. It is reasonable to have similarity values between 0 (no match) and 1 (full match). The numeric at-tributes of a project can be used easily to compute such a measure. Even for symbolic attributes some distance measure can be defined.

One more application of browsing oriented retrieval is the exploration of tools able to process a given resource — e.g. a translation memory or a term-bank — or to find the available resources that can be used when operating a given tool — e.g. a dictionaries used by a machine translation system (see figure 1).

4. Agent and resource selection

The first step of route construction is the selection of agents and resources, which are relevant with respect to the project profile. TransRouter is capable of handling several kinds of selection rules, which will be described in this section.

Type specific rules do not apply to individual objects but to object classes. Such a rule may express the fact that machine translation systems in general are of no use for projects with certain features. Currently the following type specific rules are built into the system:

- If there is no previous version of a project and no further version is to be expected and if the repetitiveness of the text is below a certain threshold, then no translation memory should be built.
- A machine translation system should not be used if the complexity of the text exceeds a certain threshold.

Most of the selection rules implemented within TransRouter apply to individual objects.

- The agent or resource must support the language pair required by the project.

- Human agents may act in different roles (translator, reviser etc), which represent individual translation services. The translator must know the required language pair (target language in case of revising) and be able to provide one of these services at the required quality level.

- Dictionaries and term banks must cover the same content domain as the project's text.

- A translation memory must be derived from a prior version of the project. This will prevent the system from doing costly assessments on memories, which probably will not have a sufficient coverage.

- Machine translation system must be able to provide the required quality.

- An agent must be at hand that can process the resource being in question.

Weak selection rules cover phenomena which are mere obstacles in the use of a resource or system. Examples are licences being outdated or format not matching. These obstacles will lead to the exclusion of a system or resource if and only if an alternative is at hand. Otherwise the rule will be suppressed in order to get some operational routes. A comment on the problems with these systems / resources is provided.

Licence must be up to date. If no other tool is available a licence can be updated easily.

- Tools must be able to process the storage format of the project's text. In most cases it should be possible to convert the formats with some reasonable effort.

- Tools must be able to produce the desired destination storage format (see above).

- Resources must be approved by a translation manager.

- Human translators must be knowledgeable in the content domain of the text and know about the relevant text styles. If nobody is at hand who has this knowledge, somebody knowing the languages should be able to do at least a low quality translation. A good reviser can sort out quality problems in the last step of the route.

After the completion of the automatic selection process the translation manager may want to reduce the relevance set further. He may know that some translator is occupied by other projects or for some reason does not want to use a specific tool etc. Removing agents or resources at this stage of the decision process will make the task of route construction and assessment less complicated.

5. Route construction

The explanation of the route construction process first needs some introduction to the route model of TransRouter. Then it can be shown how a route will be furnished with agents and resources.

5.1 The route model of TransRouter

The route model of TransRouter is comparatively simple. A route basically consists of three processing steps. Each step is performed by one main agent using a set of tools operating on a set of resources associated with this step. The pre-processing step covers all activities, which are necessary to prepare text and or resources — initial proofreading, enhancing dictionaries. The main step contains the translation process whereas the post-processing step deals with all activities following the translation until the end of the project — e.g. proofreading, formatting. Pre- and post-processing steps are performed by humans. The main agent of the translation step may be a tool as well (e.g. a machine translation system). According to the nature of the main agent the type of translation step and route will be defined. Because each type of main agent has its own requirements regarding pre- and post-processing, there are special subtypes for these steps too. The route type will have consequences on the time, cost and quality estimations as well. The agent of a translation step can make use of tools — e.g. a terminology management system or a translation memory system — which themselves will operate on data derived from special resources — term banks or translation memories. Currently Trans-Router supports the following route types:

- Translation by a translator who is employed by the agency

- Translation by a freelancer

- Translation by a machine translation system

- Translation by a translation memory system (automatic mode)

Human translators can be assigned to various roles in the translation process. They can be the main agents of the translation step if the respective sub-profile covering the translation performance is filled out. Another role would be that of a reviser or post-editor — main agents for the post-processing step — requiring further sub-profiles.

5.1 Generation of routes, assigning tools and resources to routes

TransRouter has some basic understanding of which kinds of agents and re-sources can be combined and which kinds of route steps they may be assigned to. The system will not try to generate and assess all possible combinations of main agents, tools and resources but to find some reasonable equipment for each step. This process starts with the main translation step. The system will select resources first, because the content of a term bank or a translation memory is assumed to be

prior to the effects of handling the software. In a following step the system will find the optimal tool for each of the selected resources. Finally those tools which do not need any resources (e.g. an alignment tool) will be assigned to a translation step. This approach has two implications. The system must provide the means to find a ranking of resources and tools in order to find the best fit. The solution found may not be the optimal one, because a slightly less optimal resource may be processed by some more user friendly or efficient tool, which could not be used for the resource selected. The equipment of the auxiliary steps will follow almost the same procedure with the only exception that, if possible, the same resources and tools will be used as assigned to the main step.

Not every combination of translation steps and agents is possible. A selection of rules applying is listed here:

- The agents of pre- and post-processing steps are human translators.

- The step type defines the agent type of the main step.

- If the main agent of the translation step is a tool, the main agents of the pre- and post-processing steps must be experienced in using the tool.

- The profile of a human translator must indicate that he may take the appropriate role — translator, reviser, pre- or post-editor — in the route step. This means that a processing performance greater than 0 must be assigned to this specific activity.

The most straightforward approach to the sorting of resources implies the use of resource assessments.

- Translation memories are sorted according to the coverage of the project's text.

- The sorting of term banks makes use of information on the number of unknown terms within the text.

Unfortunately the assessment of resources is time consuming. Therefore it can't be assumed that all resources which are of some relevance to the project are assessed. Thus, TransRouter has to employ two sorting strategies. If all resources of some type are assessed, TransRouter will use the assessments for sorting. If this is not the case, TransRouter has to use an alternative strategy using basic resource features for sorting.

- Translation memories are sorted according to their position in the version hierarchy. The translation memory, which is most recent with respect to the current project, will be preferred. It is most likely that this memory will have the best coverage.

- A good indicator to estimate the quality of a term bank is its size. This

largest term bank probably will have the best terminology coverage.

The sorting of agents imposes fewer problems than the sorting of resources. Basic features of the respective profiles may be used.

- Translating (translators, machine translation systems) agents are sorted according to translation quality and performance.

- The relevant features of freelancers are quality and costs.

- Translation memory systems judged according to their performance (average access and storage time).

Nevertheless, some of these data can be fully estimated only with knowledge of the complete route data. The performance of a translator for instance depends on the tools at hand, the quality of a machine translation is related to the quality of the resource being used. TransRouter will feed as much information into the sorting and ranking process as is available within the current state of decision making. In an early step only the agent profile will be avail-able, in a next one a project profile will be added. Finally all data of the route and route steps currently being elaborated are available and can be used for agent assessment.

5.2 Dependencies between agents

The sorting process as described above does not take any dependencies between agents into account. Nevertheless it seems to be quite obvious that a terminology management system which is an integral part of some other tool being used (machine translation system, translation memory system) is to be preferred to others which are not. The same applies to alignment tools or even translation memory systems. TransRouter distinguishes three levels of integration (built in, add on, compatible output). Human agents or freelancers on the other hand are more experienced in the use of some tools compared to others. These dependencies are represented in the agent's profiles and will be used in the construction of routes as follows:

- If the main agent is human, TransRouter will prefer tools that are familiar to the translator. Furthermore it will prefer tools which are able to mutually cooperate. The level of integration will be considered only if there is no severe lack of performance compared to some other tools.

- If the main agent is a tool — e.g. a machine translation system — it is re-quested that all tools assigned to the main translation step allow some integration with the main agent.

5.3 Manual modification of routes

The system, as already has been mentioned, will not necessarily find the optimal route. But even an optimal route could be of little use, if the agents of the route

were occupied with other projects. In this case there is a need for the translation manager to modify routes suggested by TransRouter manually. He may delete entire routes or copy routes to try out different versions of the same general approach. Possible modifications of a route include the re-placement of the main agent of a step, the removal of tools or resources from a step or the assignment of additional or alternative ones. This process is governed by a set of simple rules.

- The main agent may only be replaced by an agent of the same type. The route type will be unaffected. Changing the main agent will trigger a consistency check on tools and resources assigned to that step. It is checked whether the new agent may use them. If this is not the case the resource or tool will be replaced as well.

- If an agent is removed from a step, the corresponding resource will be removed too and vice versa. This will prevent the user from constructing inconsistent translation steps containing useless tools or resources lacking an agent.

- If a new resource is added to a step, TransRouter will remove an equivalent resource (same type) from the route if present. If the agent corresponding to the replaced resource is not able to process the new resource it will be replaced too. The optimal tool, which can process the needed data format, will be chosen automatically. An equivalent process will take place if an agent is replaced. Since all steps of a route should have the same equipment if possible, these exchange processes are performed on all steps simultaneously if the new agents or re-sources are valid for all of them. Otherwise the manipulation is restricted to the explicit manipulation of a single step.

6. Route assessment

The step of route construction is followed by the assessment of routes. This step will be performed automatically. The user can guide this process only by adjusting the criteria which have to be used. Generally the assessment of routes can serve different purposes (see also figure 3):

- Ranking of routes will help to find the best route with respect to a set of criteria.

- Estimation of time, cost or quality numbers will be helpful for the final planning processes. Setting a frame for time, cost and quality is an important task at the very start of a project.

In the course of the workflow supported by TransRouter the ranking of routes would be the first step. An exact estimation could be restricted to the best routes one should consider implementing. Even from the viewpoint of the de-signer of a decision support system this sequence seems to be reasonable. Where the

ranking of routes is a rather feasible task, the exact estimation of route features imposes severe methodological problems:

- Some of the relevant criteria — this is true especially for quality — are not well defined.

- The nature of translation processes up to now is not well understood. The effects of the environment — features of projects, agents and re-sources — on the translation process with respect to time, costs and quality can — in many cases — be quantified only by very rough approximations.

- There are aspects of the handling of a translation project, which are idiosyncratic to any translation agency.

- Some cost relevant issues can be discussed only on a larger scale than a single project. What is the benefit of a high quality translation memory? Which share of a software licence has to be charged for?

- An exact estimation of cost, time and quality would require rather de-tailed data about projects, agents and resources. It is questionable whether the result would justify the effort of data acquisition.

As a consequence TransRouter will provide mechanisms for the ranking of projects according to cost, time and quality. Additionally means for the computation of total values are provided. These depend on domain specific knowledge that cannot be provided in the scope of the project. Means for the acquisition of these data, however, are present.

A full description of the route assessment mechanisms of TransRouter would exceed the scope of this paper. As a consequence the general mechanism will be explained at the example of time assessment. Some remarks on quality will follow. Cost estimation will not be dealt with at all.

6.1 Translation time

The most basic understanding of translation time can be defined in a few sentences.

- The time needed to process a translation route is the sum of the processing time of all of its steps.

- The translation time of a route step is computed from the number of words of the text times the agent's (translator, reviser etc) performance as contained in his profile (measured in words per hour).

Obviously this formula is a crude abstraction because there is no single translation performance of a translator. Performance in translation is dependent on a number of factors the most important of which will be named here:

- The first important factor seems to be the language pair. A translator may be competent in several languages but the translation performance will vary.

- High translation speed will have a negative impact on quality. Thus translation performance will decrease in projects with higher quality requirements.

- Subject domain and text style will also be influential. The knowledge of a specialised vocabulary or of conventional rules of text structure and formulation might be necessary.
- If the text contains a great deal of new terminology, this will affect translation performance adversely.

- Certain text types, for example legal text containing quotations which must be quoted rather than re-translated, will have an adverse effect on translation performance (because of the time required to search out the quotations): mitigated of course, if an appropriate translation memory is at available.

- Translation performance probably will depend on the readability or complexity of the text. TransRouter provides a tool for the estimation of text complexity. Since the notion of text complexity is not well understood up to now, this estimate can only be heuristic.

The influence of these factors seems to differ between individual translators. Thus, an exact translation performance measure would require the empirical acquisition of a huge matrix of interdependencies. Since this is not feasible a sufficient approximation must be found. The level of detail, which is required or asked for, will differ between translation agencies. Thus, TransRouter will support the definition of rather general and quite detailed models as well.

The mutual dependencies between translation performance and project features are of major importance for the time estimates of TransRouter. Trans-Router uses an associative access method based on keys of variable length (see also figure 2). This mechanism will be described on the base of performance mapping as an example. The same mechanism will be used for other complex features — translation quality and translation costs — as well.

Definition of absolute translation speed

Each profile of a translator or machine translation system contains mappings from project features to translation performance values (speed, quality, cost). The current implementation uses all relevant project features (source and tar-get language, subject domain, text style and complexity, formats etc.) for keyed performance access. Additionally information about the translation route may be used (tools being used, features of resources being used). To avoid data acquisition overhead, a partial definition of access keys is possible. A fully

unspecified key will retrieve a default value from the system. An access key may contain the following `wildcards' instead of true project or re-source data:

- '*': Matches any value. This is useful especially if a default value is to be defined that is valid for any project constellation.

- 'some': Matches any value other than the empty object. This is useful for instance if a default value for pre-editing for machine translation is to be defined. In this case at least some machine translation system must be present within the route.

- 'none': Matches if only the empty object is present. This is useful if the use of specific system or agent types should be excluded.

Using this specification method the following phenomena can be expressed easily:

- A machine translation system can handle the following six language pairs at an average speed with given quality. The language pairs will be defined in the profile. No specific performance keys will be used. System performance will be defined as a default.

- The system will translate English to German at a higher speed. A specific key *(source:* English, *target:* German) concerning this language pair will be entered.

- A translator is responsible only for some very specific cases (e.g. scientific reports about biology). A specific key covers the respective translation performance. The default translation performance will be set to zero, thus prohibiting the assignment of other projects.

To access performance data for a project or route the following steps will be performed:

- Derive an access key from the project's (route's) features.

- Sort the access keys of the profile according to the number of features specified in descending order.

- Select those keys which subsume the access-key for the chosen project or route.

- Map the selected keys to their performance values.

- Compute a singe value from the selected ones. In the case of translation speed and translation quality this means using the smallest value. Translation speed, thus, is defined by upper bounds, which are valid for specific situations. In case of cost computation every key will contribute a cost factor adding to the total costs.

Definition of factors affecting translation performance - attached procedures

Obviously this matching algorithm, which prefers the most specific access keys, does not allow the use of general rules. More general effects are not described by absolute values but by numeric factors or even attached procedures. TransRouter will use a specific key structure (subject domain, text structure, text complexity) to access these data using the following algorithm:

- Derive an access key from the project's features.

- Find all keys matching the search key.

- Compute the product of all factors, which are associated to these access keys.

- Compute product of the resulting cumulative factor and the already known total performance value.

- Additionally attached procedures may be retrieved. These code fragments are sorted according to an inherent precedence value and then arranged as a pipe. The performance value found so far is used as input to this pipe. The output of the pipe is the final performance value, which will be further used. The attached procedures accept three parameters: 1st is the translation step currently being elaborated (or nil), 2nd is the project profile and 3rd is the translation performance value that has been computed so far.
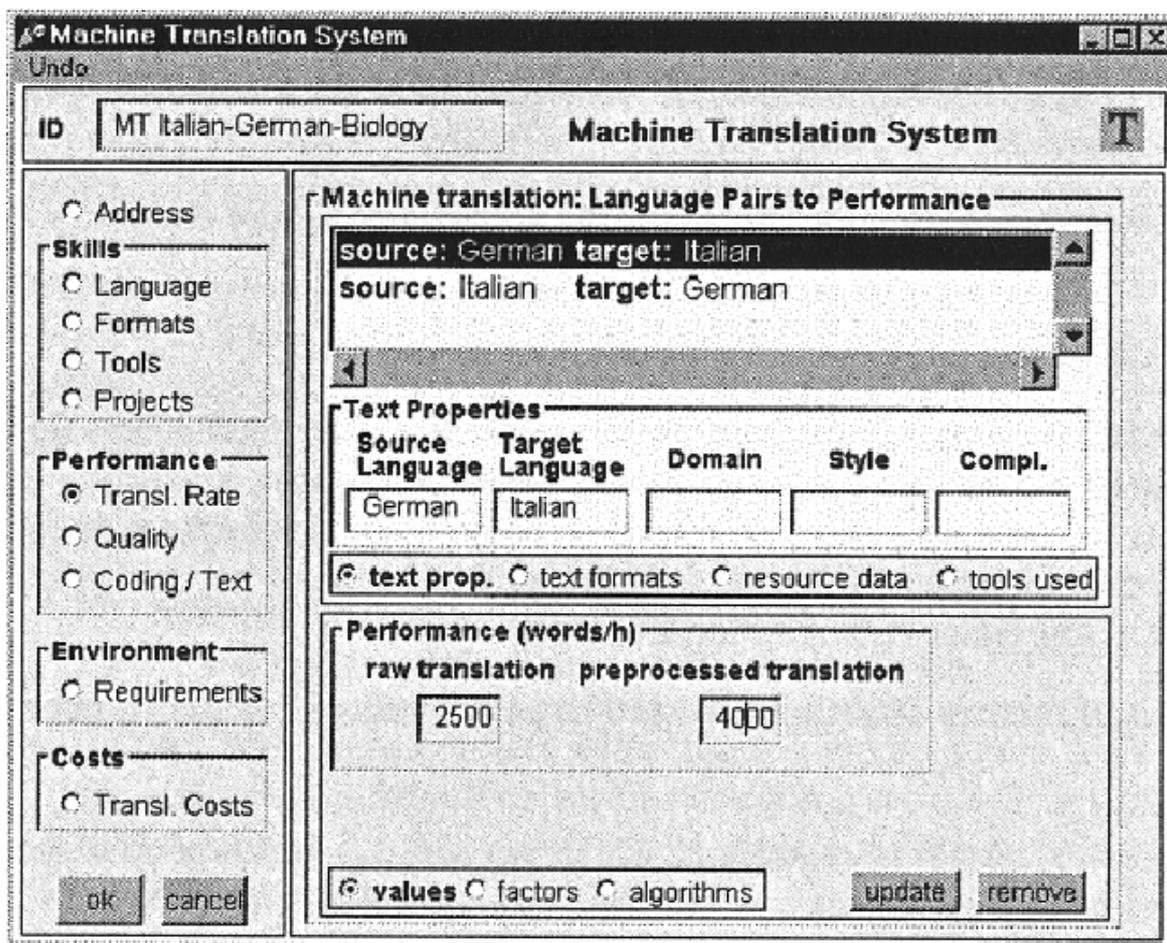
Figure 2: Profile of a machine translation system defining the mapping of project data to performance values

The expressive power of TransRouter now is extended to phenomena like the following:

- A translator translates scientific texts by an excess of 30% of average translation time. Assign a factor of 1.3 to scientific texts.

- The use of a specific tool increases productivity by 15%.

- Do not even consider using a specific machine translation system to translate texts of legislation. Assign a performance factor of 0 to a general key *(domain:* legislation).

- The translation performance will not exceed a certain threshold, if the text is very complex. In this case a code fragment will check the threshold.

- The performance of teams is computed by some algorithm, which is defined as default (see below). As a consequence this algorithm can be modified easily according to the specific needs of an organisation.

Defaults

A final extension of the data model allows the definition of default parameters for agent classes. A default performance profile for each type of agent concerned with translation (translators, freelancers, and machine translation systems) and the related activities (revising, pre- and post-editing) is available. In a commercial environment a system like TransRouter would probably be de-livered with agent profiles (except translators) and defaults being set.

Default profiles have the same structure as those of individual agents. Thus, the same phenomena can be expressed. Nevertheless, default profiles will contain only a few absolute values but most of the general factors and attached procedures of the system. Individual profiles on the other hand will contain absolute values, which will be modified by factors or procedures de-rived from the default profiles.

Default profile and individual profile will be merged on access time. Each key and value pair of the default will be moved to the individual profile if and only if there is not a similar key already in existence in the individual profile. Thus, definitions in individual profiles take precedence over those of default profiles.

Preinstalled defaults

Some defaults concerning attached procedures will be defined automatically by the system at installation time. They may be modified later on according to the specific needs of an organisation or user.

- Team performance: If a translator is member of a team, his translation performance will be reduced by a certain amount to cover organisational overhead.

- The performance of a post-editor depends on the difference between the quality value of the main translation step and the degree of quality expected from the project.

Stepwise refinement of access keys

The flexible size of access keys does not only allow the choice of an adequate level of detail in the definition of data but also the stepwise refinement of access within the decision process. In the beginning only project data are avail-able. Later on additional information about possible routes and their resource assignment is at hand. Especially knowledge about the tools being used within a translation step can influence the translation performance and therefore will be included into the key structure. As a consequence TransRouter's estimate of translation performance (and quality) will be improved when the user enters additional information — especially about routes.

Translation quality

Quality' is one of the most problematic concepts within translation evaluation [Marx et al 98]. There is neither a clear definition of the concept of text or translation quality nor a sufficient understanding of the interrelation between the translation process and its outcomes.
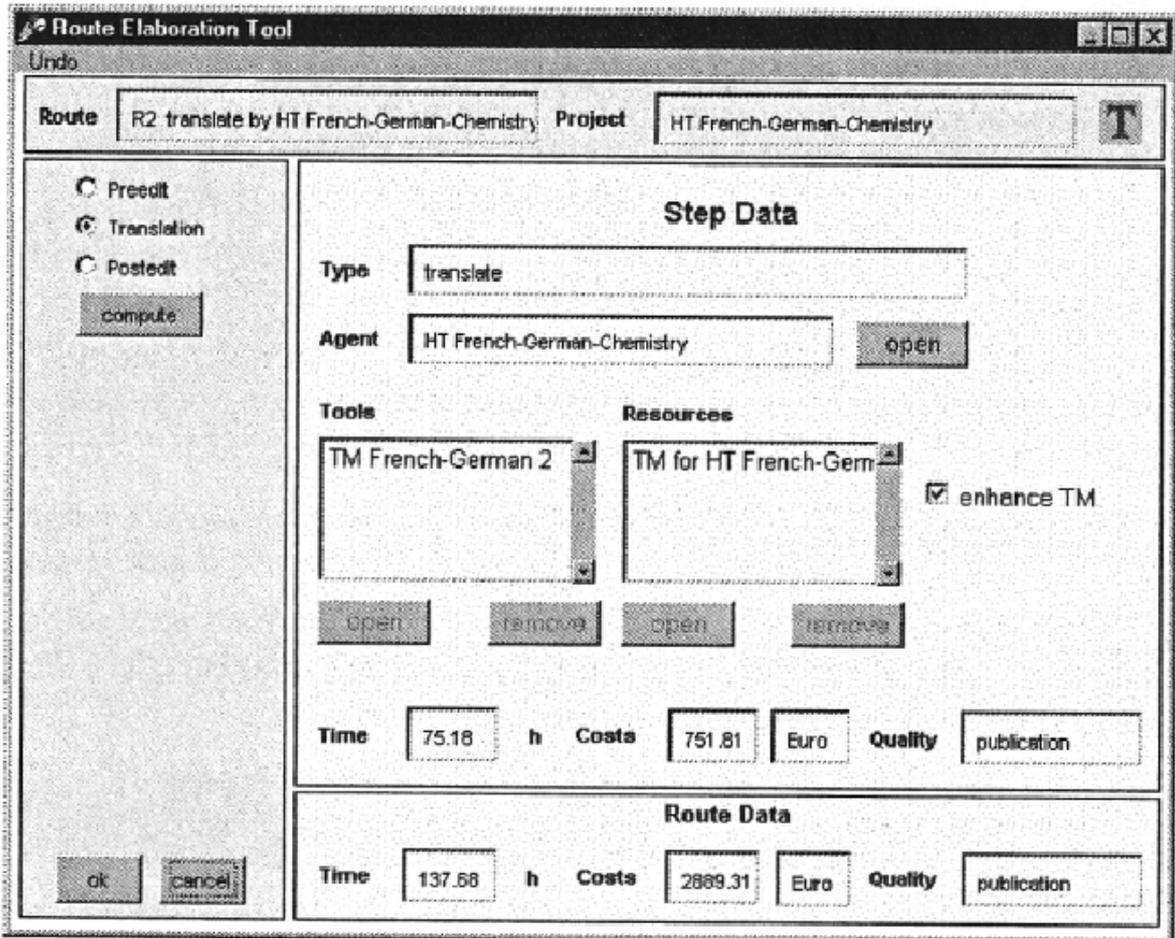
Figure 3: Assessment of a translation route with emphasis on the translation step

Within a conventional production process quality is described as the probability that an individual product has the required features. A clear definition of product features, which is usually given in a product description (design, modes of operation etc.), is a prerequisite of this approach. Additionally a sufficient number of similar objects must be produced in order to be able to compute probabilities. Neither of these conditions is fulfilled in the case of text translation. Every translation is a very individual product presumably not allowing the estimation of fault probabilities. Most of the quality criteria that can be agreed on can not be formalised in a way that it can be used by a decision support system.

The lack of an exact quality measure is a common problem for service providers. One solution of this problem is to define quality not primarily as a feature of a

product but as a feature of the process of product construction or service delivery. From this point of view a tool like TransRouter is a major means of quality management since TransRouter will show possible translation routes and name the quality effects that can be expected. Quality requirements can be defined with respect to specific translation tasks [White/Taylor 98, Pavlsen et al 98]. This is a major step forward even if a general quantitative model of quality can not be provided.

Quality values are communicated to the user by four symbolic values:

- No use: The translation will probably be in such a bad shape that it can not be used at all.

- Browsing quality: The reader will be able to identify what the text is about.

- Information dissemination quality: The reader will identify the arguments and major facts of the text.

- Publication quality: The translation fully meets the quality standards of the original version with respect to content and form.

The quality of a translation will primarily result from the competence of the translator or translation system and the quality effects of translation tools and resources. Thus, TransRouter allows the definition of a detailed quality profile for translation agents. A quality profile generally is a mapping from project data to quality values. The access method is the same as described earlier in this paper. Similarly quality effects can be described using absolute values, factors or algorithms. Similar to translation performance all relevant project and route features — language pair, domain, text style, TM coverage, unknown terminology etc. - are covered.

7. Project documentation and learning from data

Since TransRouter is a tool for supporting a translation manager in the appropriate choice of a route for a translation project. Issues concerning the project management and documentation seem to be outside the scope of the project. This is only true from the point of view of a single project. The decisions taken in a project and the outcome of these decisions however constitute valuable information for a new project if the project features are comparable in some way.

This is especially true as most of the data initially fed into TransRouter can only by approximations or even guesses. If projects that are processed using TransRouter are documented, the translation manager can get a notion of the quality and value of the decisions taken by TransRouter.

Thus, TransRouter will ask the translation manager for the following information — if not already present - on every project and put it into the archive:

66

- The project profile

- The route that is finally chosen after the decision process, including the agent, tool and resource assignments and the calculated time, cost and quality values

- The route as it is implemented within the course of the translation project, including the agents, tools and resources that were finally assigned to the project and the time that was needed for each translation step and the quality which was achieved.

On the basis of this information TransRouter can assess its own performance and calculate the averages of the deviation between predicted and true values for all projects and — more decisive — for all projects similar to a new project which is to be tackled by TransRouter. Furthermore all tools or resources may be identified, which although chosen by TransRouter on the basis of their pro-files are often skipped by the managers in the real implementation of the project. These simple but useful features are already implemented in the Trans-Router prototype.

The next step in the evolution of TransRouter, which cannot be taken within the scope of this project, would be the learning from agent features from real world data. Every translation project represents a new case from which system parameters can be learned. The first step would be the acquisition of very specific access keys to performance and quality that represent the relevant project features. These keys will then be assigned to the agent profiles. If similar cases occur later on, the data may be adapted to get a best fit to all similar cases. Later on, when a fair sized pool of cases is at hand, generalisation processes may be run on the profiles. They will isolate those project and route features, which contribute significantly to the project outcome and skip those that don't. Thus, the general predictive quality of the system will be gradually enhanced.

References

James R. Evans and William R. Lindsay: The management and control of quality. South-Western College Publishing, 1999.

Entscheidungsunterstützung bei der Planung von Übersetzungsprojekten. In: Zimmermann / Schramm (eds.), Knowledge Management and Communication Systems, Proc. of the 6th Int. Symposium on Information Science, pp. 47-57, 1998 Hammwöhner 00 Decision support in the routing of translation projects. Re-port on the Regensburg prototype of TransRouter. Technical Report, 1999.

The TransRouter consortium (contact author: Margaret King): TransRouter: a decision support tool for translation managers. In. Proc. MT Summit Conference, September 13-17 1999, Singapore, 1999

Jutta Marx, Nancy Smith and Bernhard Staudinger: Some Problems in the

Evaluation of the Russian-German Machine Translation System MIROSLAV. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol. 2, pp. 1219-1226, 1998.

Claus Pavlsen, Nancy Underwood, Bradley Music and Anne Neville: Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol.1, pp. 27- 31, 1998.

Efraim Truban: Decision Support and Expert Systems. Management Support Systems. Prentice Hall, 1995.

John S. White and Kathryn B. Taylor: A Task-Oriented Evaluation Metric for Machine Translation. In: Rubio et al (eds.), Proc. of the First Int. Conf. On Language Resources & Evaluation, Granada, 1998, Vol. 1, pp. 21- 25, 1998.