

The IAMT Certification initiative and system categorization

John Hutchins

The need for guidance

- multiplicity of new systems every month
- new types of systems
- new customers, purchasers, users
- greater exposure to information in foreign languages
- widespread misconceptions about nature of translation
- high expectations about computers, and speed of software development
- unawareness of the complexity of language
- ignorance of reasons for slow progress in MT

‘Problem’ of the name

- “Machine translation” is seen as an old-fashioned term; and it does not cover computer aids for translation.
- “Translation software” is reasonable for covering commercial products, but is not relevant for research and development activity.
- “CAT (Computer-aided translation)”, does not cover fully automatic systems
- Other terms: “Translation technology”, “Computational translation”
- But “MT” has ‘merit’ of long established usage and some familiarity outside the MT community

What is needed

- Guidance to what is available
- Guidance on suitability of different types of systems for particular uses and tasks
- Information on components of systems and how they may be used (inc.benefits and limitations)
- Information on cost-effective and/or time-saving use of different types of translation support and MT systems
- Guidance on evaluation of systems
- Establishment of benchmarks

IAMT Certification Group

- Set up in 1997 under leadership of Eduard Hovy (AMTA president)
- Panel at AMTA 1998 conference
- Informal group proposing definitions of system types
- Aim: some kind of ‘official’ IAMT certification to aid vendors and potential purchasers
- Preliminary definitions circulated during 1999

“Compendium of translation software”

- Parallel independent effort
- A guide to commercially available systems
- Includes an ‘unofficial’ categorisation of system types
- Also some guidance on components

Certification: aims

- Categorisation of systems:
 - easily understood by potential users
 - for assisting selection by purchasers
 - agreed and adopted by vendors
 - perhaps with symbol on products
 - basis for comparative evaluations (consumer reports, company evaluations, etc.)
 - basis for future benchmarking

Basic distinctions (for general public)

- wholly automatic systems
 - systems that (attempt to) translate texts and sentences as wholes
- computer-based translation aids
 - systems that provide linguistic aids for translation: dictionaries, grammars, concordances, etc.

Types of aids and support systems (1)

- Dictionaries: both bilingual and multilingual, with and without grammatical information, with and without guidance on usage (appropriateness)
- Language aids: grammatical information (morphology, noun/verb paradigms)
- Spelling checkers
- Style checkers
- Terminology aids: glossaries of ‘authorized’ terminology

Types of aids and support systems (2)

- Specialised glossaries: for translator's special subject areas, for clients, agencies and customers
- Input controls: pre-editing tools, controlled language
- Corpora of 'approved' translations (translation memories and alignment tools)
- Management support tools: project control, budgeting and cost controls, workflow and personnel management, etc.

Minimal basic features of MT

- More than word for word
- At least minimally correct morphology (endings of adjectives agreeing with nouns, noun cases agreeing with verb forms, etc.)
- Some minimal syntactic processing (e.g. inversion of adjective and noun when translating from English to French, movement of verb position when translating from German to English)
- Some semantic processing (e.g. selection among alternative ‘equivalents’ according to context or subject field)
- Input by phrase or sentence (not word by word)
- Output in ‘complete’ sentences

Defining MT

- For the purpose of categorization (for general public) the definitions should not be given in terms of internal processes, e.g.
 - interlingua, transfer, example-based, statistical analysis, feedback/learning, ‘artificial intelligence’.
- These are relevant when discussing research systems but not for the description of commercial systems.

Sources of confusion

- Some vendors refer to electronic dictionaries as ‘translation systems’ since capable of ‘word-for-word’ renditions
- Some MT systems are little more than dictionaries
- Terms ‘computer-aided translation’ (CAT) and ‘machine-aided translation’ used for both HAMT (human-aided MT, e.g. post-edited MT) and MAHT (computer-based translation tools)
- Some systems combine MT and translation support tools
- Translation memories can be either translation aids or components of MT systems (some MT systems based on TM technology)
- MT systems can be components of translator workstations

Definitions of basic categories: MT system

- *(IAMT Certification group)*
 - A software system is a "machine translation system" if it takes input in the form of full sentences at a time and generates corresponding full sentences (not necessarily of good quality).
- *(Compendium)*
 - MT system: software for automatic translation, where input units are full sentences of one natural language and the output units are corresponding full sentences of another language.

Definitions of basic categories:

Translation support tool

- *(IAMT Certification group)*
 - A software system is a "translation support tool" if it takes input one word or phrase at a time and produces proposed translated equivalents, which users must then assemble into sentences and texts.
- *(Compendium)*
 - [No general definition provided. Instead types are listed]

Spread of usage levels

for both MT systems and translation support tools

- Basic level [or ‘entry’ level, ‘home use’]
 - Systems not considered suitable for professional translators and not expected to produce top-quality translations; also relatively cheap
- Standard or Professional level
 - Systems which may be found cost-effective for the professional user, while not expected to produce translations of ‘publishable’ quality without revision.
- Company level
 - Systems intended for large-scale processing (repetitive documents, multiple languages, technical documentation) and which may be integrated into other documentation processes (authoring, controlled languages, publication, etc.)

Levels of MT systems

IAMT Certification Group

- Machine Translation System (basic level)
 - A machine translation system is a "Basic level system" if its largest dictionary contains fewer than 50,000 entries OR if has very limited facilities for users to extend the dictionary OR if its translation capability is restricted to the translation of single-clause (basic) sentences. Systems at this level are primarily meant for home use.

Levels of MT systems

IAMT Certification Group

- Machine Translation System (standard level)
 - A machine translation system is a "Standard level system" if it contains a dictionary of more than 50,000 root entries AND it is not restricted to the translation of single-clause (basic) sentences AND it provides facilities for the user to extend the dictionary. Systems at this level are primarily meant for home use or standalone office use.

Levels of MT systems

IAMT Certification Group

- Machine Translation System (advanced level)
 - A machine translation system is an "Advanced level system" if its smallest dictionary contains more than 75,000 root entries AND it is not restricted to the translation of single-clause sentences AND it provides facilities for the user to extend the dictionary AND it supports a configuration in which more than one client can be networked with a single central server computer. Systems at this level are primarily meant for office use with networking capabilities.

Problems with definitions

- Specification of the various components:
- How to achieve concision, neutrality and avoid constant changes
- Dictionary sizes: how defined (entries, stems, translations, etc.); general, special and user dictionaries
 - (NB. user dictionaries now feature in most systems)
- Networking (not restricted to ‘advanced’ systems)

Definitions in “Compendium”

- Intended to be oriented towards types of user, types of use and tasks
- Conform to some extent with vendors’ own descriptions
- Aim to avoid indications of translation (output) quality
- Entries for particular systems provide additional information intended to assist the selection of systems (e.g. hardware and software requirements, dictionaries, translation speeds, and prices).
Components are not included in definitions
- The categories are general indicators of systems types, defined in terms of what users might expect from such systems.

“Compendium” definitions of fully automatic systems (1)

- MT system: software for automatic translation, where input units are full sentences of one natural language and the output units are corresponding full sentences of another language.
- MT system (home use): an automatic translation system designed by the producer for personal use by the general public, i.e. by persons not normally with any experience or training in translation. (Caution: some of these systems are little more than dictionaries.)

“Compendium” definitions of fully automatic systems (2)

- MT system (for Internet/Web): system developed specifically for translating electronic documents on the Internet such as electronic mail, Web pages, chat discussions, etc.
- MT system (professional use): system designed for use by professional translators, usually working independently or for translation agencies
- MT system (client/server): system designed for company intranets to support a team of translators (often professionally trained)

Definitions of translation support tools

(1): Electronic dictionaries

- (*IAMT group*)
 - A translation support tool is an “Electronic Dictionary” if it consists mainly of a bilingual or multilingual dictionary together with a dictionary lookup facility or mechanism.
- (*Compendium*)
 - Electronic dictionary: bilingual or multilingual database of lexical entries (words or phrases) searchable individually or in combination, either for consultation or for insertion into human-produced translations.
- In the Compendium there is also a note: (A dictionary might also be used to produce ‘word-for-word’ renditions, i.e. sequences of individually translated words in the syntactic order of the original texts.)

Definitions of translation support tools

(2) Terminology management systems

- *(IAMT group)*
 - A translation support tool is a “Terminology Management System” if it consists mainly of methods to help the user construct a multilingual terminology dictionary, together with a dictionary lookup facility implemented as a database (and not simply as a word-list), and includes multiple fields utilized according to standard terminology practice.
- *(Compendium)*
 - Terminology management system: software for the creation, maintenance and searching of multilingual databases of terminology compiled for local (company or personal) use.

Definitions of translation support tools

(3): Translation memory system

- (*IAMT group*)
 - A translation support tool is a “Translation Memory System” if it consists mainly of methods to help the user construct and use a collection of sets of previously translated text (ranging in length from phrases to whole texts) together with one or more of a text alignment facility, a text lookup facility, and a storage management facility.
- (*Compendium*)
 - Translation memory system: software (or component of system) for the creation, maintenance and searching of bilingual databases of previously translated texts.
- [NB Compendium has ‘Alignment tools’ as separate category]

Definitions of translation support tools

(4): Foreign language authoring tools

- *(IAMT group)*
 - A translation support tool is a “Foreign Language Authoring System” if it consists of lexicons, phrases, and even full text examples that assist the user to write documents such as business letters, contracts, etc., in a language in which they are not fluent.
- *(Compendium)*
 - Foreign language authoring system: software enabling composition of texts (e.g. business correspondence or electronic mail) in another language not necessarily from an original source text.

Definitions of translation support tools

(5): Translator workstation

- *(IAMT group)*
 - A translation support tool is a “Translator Workstation” if it consists of several Translation Support Tools integrated into a single framework.
- *(Compendium)*
 - Translator workstation: integrated system for the use of professional translators, which combines (normally) multilingual word-processing, terminology management, translation memory, and (optional) automatic translation.

Additional definitions (“Compendium”)

- Localization support tool: system for the cultural and linguistic adaptation of products (computer software) and their documentation
- Alignment tool: software for the creation of bilingual text databases where sentences (or phrases) of source texts are linked to corresponding text segments of a target language.
- Pre-editing tool: software for the preparation of input texts, often including means for the control of input language, i.e. the reduction of ambiguities and the simplification of structures in order to facilitate automatic translation.
- MT service: translation service via Internet (or mobile telephone), using MT systems with or without human post-editing, and charging according to length and/or subject of texts. (Some on-line Internet services are free.)

Alternative approach: categorization by components

- Instead of defining system types, perhaps define components/facilities of systems (both MT and translation tools)
- System categories would then be defined secondarily as collections of features/facilities.

Features/facilities (tentative)

(C= Company MT, P=Professional MT, H=Home MT, T=Translator workstation)

- Pre-editing C, P
- Controlled input C
- Dictionaries (general) C, P, H
- Special dictionaries C, P(?)
- User dictionary(ies) C, P
- Terminology management C, P, T
- Post-editing C, T
- Concordancing P
- Translation memory C, T
- Alignment tool T
- Document formatting, conversion C, T
- Publishing C, T
- Web page translation C, H
- Internet/on-line access C, P, H, T

Features and systems: towards alternative definitions

- Systems defined in terms of facilities and features, e.g.
 - ‘Company use’ MT system: facilities for pre-editing, post-editing, terminology management, networking, user dictionaries, etc.
 - ‘Home use’ MT system: general dictionary, Webpage translation, etc.
 - Translator workstation: user dictionary, terminology management, translation memory, alignment, etc.

Evaluation and benchmarking: requirements

- Build on certification/categorisation process
 - if possible, all measures should be applicable to full range of systems (MT systems and translation support tools)
- Evaluations of systems as wholes and of their components:
 - Comparisons of systems (of same type or category)
 - Comparisons of features/facilities (of different systems, even of systems not in the same categories)
- Benchmarks:
 - Establishment of minimal sets of facilities for each system type
 - Establishment of minimal levels of performance for each type of facility (within each system category)

Evaluating systems for quality

- Traditional approach
 - assumed/expected by most of those wanting evaluations
- Evaluation of systems by *qualitative* measures, e.g.:
 - faithfulness of output to original sense
 - intelligibility of output (independently of original)
 - accuracy of terminology
 - readability (as if in original language, not translation)
 - stylistic appropriateness (language, subject, readership)

Evaluating systems for utility

- Evaluations of *utility* or *usability* might include:
 - Cost saving (time, money)
 - Ease of use (by experts, by novices)
 - Level of intelligibility, accuracy, etc. appropriate for expected use
 - Speed, response times
 - Effect on organisation's overall translation throughput
 - Compatibility with other systems (authoring, publishing, terminology, etc.)
 - Relative importance of quality and speed (for particular user group)
 - [NB. Evaluation of usability may sometimes be more valuable than evaluations of translation quality]

Evaluation of features

- Each component feature/facility (of each system) evaluated in terms of quality and/or usability
- Evaluation measures according to specific user or user group
 - what they like about each feature
 - what features they would like to be included
 - how they rank features in importance/value

Benchmarking

- Minimal levels:
 - presence of components/facilities (in products of same system type)
- Performance measures:
 - initially: benchmarks for facilities and components
 - eventually: benchmarks for systems as wholes, according to performance of each component

How IAMT might help

- clear descriptions of system types, facilities and capabilities
- description of factors that companies might take into account when deciding on which systems to consider for evaluation, and when deciding how to evaluate
- description of features/facilities of different system types, for the ‘professional’ user (translator)
- description for the occasional (‘home’) user of what MT can and cannot do
- description of what translation support tools are available and how they may be used.

How IAMT might help: Evaluation and benchmarks

- Establish some agreement on evaluation criteria, both for systems and for features (components)
 - initially: system categorisation and feature specification
 - in collaboration (or co-ordination) with other relevant bodies, e.g. EAGLES, LISA, etc.
- Offer expertise (through some delegated sub-committee or group):
 - advice in setting up benchmarking tests
 - advice to other organisations producing benchmarks
- General comment: IAMT must recognise that if it does not produce guidance for evaluations and does not establish and publish benchmarks then some other (perhaps less well qualified) organisation may do it.

Disclaimer

these have been personal views with no 'official' IAMT status

Comments (email): WJHutchins@compuserve.com

Copy of article (as in May 2000) at:

<http://ourworld.compuserve.com/homepages/WJHutchins>