DOWNLOAD PDF READINGS IN MACHINE TRANSLATION

Chapter 1 : Readings in machine translation | Sergei Nirenburg - blog.quintoapp.com

*Readings in Machine Translation [Sergei Nirenburg, Harold L. Somers, Yorick A. Wilks] on blog.quintoapp.com *FREE* shipping on qualifying offers. A collection of historically significant articles on machine translation, from its beginnings through the early s. The field of machine translation (MT)â€•the automation of translation between human.*

Initially designed in the United Kingdom in the s to give blind readers direct access to inkprint materials such as books, new versions continued to be produced in the U. Reading with an optophone requires learning a tonal code: Illustration of the exploring optophone. During the twentieth century, they were characterized as translation devices by their designers and operators, at a time when transcription and translation were often synonymous. Optophone for blind readers, patented by E. Optophones have few remaining users and are no longer commercially produced, yet the technology has been remarkably influential. The apparatus itself has given rise to multiple generations of reading machines, while its components, principle of audiovisual conversion, and synaesthetic mythos have disseminated into distant provinces of engineering and the arts. After building a new version of the optophone, Zworykin and his RCA colleagues proceeded to assemble a reading machine capable of Optical Character Recognition OCR â€" the first of its kind. Scholarly treatments have mostly depicted optophones as obsolete, framing them in the narrowest terms as discrete instruments, rather than more general technologies. The labour point of view encompasses manufacturer, developer, and reader. These early adopters popularized the optophone through their public demonstrations, lending it a reputation that far outstripped actual use. Exceeding the conventions of what came to be called human factors design and usability testing, optophone reading required extensive training; moreover, the capabilities of the technology only became evident in concert with skilled reading. The embodied experiences of blind readers were distinct from those of sighted inventors and manufacturers; blind readers likewise demonstrated greater expertise about many aspects of optophone use. Most obviously, the labour and expertise of optophone reading can be examined by treating the technology and its operators as objects of historical inquiry. Pictured in Figure 4, Mary Jameson was the first celebrity optophone reader; she was both subject and author of numerous articles on the device. She used the British optophone daily and tested subsequent American models, such as the visotoner see Figure 5 , into the s. Veterans Administration with whom Jameson corresponded via braille and tape recording. We further propose that optophones can be remade or prototyped see Figures 6, 7, and 8 , such that optophonics itself becomes the mode of inquiry into machine labour and maintenance of machine work. It is especially useful when historical devices, such as reading optophones, are minimally documented, no longer available, or do not function as they once did. Prototyping the Past is an expansive research program at the University of Victoria, where two of us Chan and Sayers are located. In this instance, prototyping enables researchers to study and test the dimensions involved in past listening, reading, coding, and decoding practices â€" without any assumption of recovering specific past experiences. If archives provide testimonial fragments about individual use, then 2D to 3D translation helps scholars in the present to broadly characterize optophone reading, weigh historical controversies, and understand technical affordances. The prototype by Chan and Sayers, in turn, prompted Mills to return to the archival materials she had previously collected on the optophone, looking for concrete design contributions of early blind readers as discussed below. Like translation more generally, prototyping does not seek a straightforward, 1-to-1 equivalency, nor does it seek to replicate past devices or embodied experiences. Instead, it highlights smaller gains or losses over time and across versions. That is, it foregrounds difference and absences: Furthermore, the process calls attention to the labour of translation â€" both in the past the labour of Mary Jameson and other optophone users and now the labour of remaking â€" that typically escapes media studies scholarship. Prototype of a reading optophone glass and handle not pictured , laser-cut and powered by a Raspberry Pi and Pi camera. Video demonstrating the process of prototyping a reading optophone. Human operators facilitated the machine transcription process, preliminary to the act of reading the tonal code. The tracer contained an element called selenium, which detected contrasts between white pages and black type and then converted this pattern into streams of tones and silences. To listen, operators wore telephone receivers over their ears like

headphones, translating the tones and silences into characters or words. They could also tune an optophone with a knob and physically control the pace and location of reading. Thomas Gilman Moorhead, another early adopter, was a distinguished physician who became blind after falling on a train platform as an adult. He was later elected President of the British Medical Association. Green was one of 80 blind typists trained by The National Institute for the Blind NIB to use a new shorthand machine, and she helped Burrows prepare the manuscript for The Story of Broadcasting , a fact he publicized in the text itself: In the writing of this book it has been my great privilege to have the assistance of one â€" a Miss Mabel Green â€" who has been blind from infancy. Nearly every word here set down in print was dictated to Miss Green at normal dictation speed. The shorthand notes were made by a machine having seven keys only, one for each of the six dots upon which the Braille system is built, and one for spacing. The phonetic signs were embossed on a paper tape similar to that used for recording telegraphic messages. The transcriptions from this tape, made by means of an ordinary typewriter, have been astonishingly accurate, and would do credit to one in full possession of his powers of vision. This book is, I believe, the first to have been prepared under such conditions. Press coverage aside, most blind schools could not afford to purchase one. Moreover, local educators at the NIB and St. I am surprised to find how easy it is to know what is coming when one becomes familiar with a writer, to pass quickly over the words, and then fancy one is making wonderful progress in speed; and then, what disappointment when one takes up a fresh book! In another letter to Barr, Thomas Jameson insisted that the success of the optophone depended on four elements â€" the inventor, the manufacturer with whom he included sales , the demonstrator and teacher whom he listed together , and the repairer. She worked with Lundie on improvements to the line-changer and to a component that allowed the reading of italics. One, Michael Lloyd, was a music teacher in Birmingham who himself demonstrated its application to the reading of music. I always begin by the general statement that the optophone translates print into musical motifs, and that the nature of these depends on the shapes to be translated. I go on to say that every vertical line produces a chord â€" examples h and I; every diagonal a tune â€" examples v and w. Having to do v I may invite those listening to tell me how w will sound. If they can, I know they are grasping the idea. It is more difficult to describe curves, but I ask people to listen to the smooth passing of the notes. I then, as you do, talk about the distribution of the notes; middle range with ascenders and descenders. I like to use p and b, q and d for this. Over and over again I have found this approach awakens an interest in the optophone, and teaches them the basic sounds. With pupils familiar with letter-shapes it helps them to memorise the corresponding motifs. She used it to read books, food labels, and typed mail. In the s and s, she corresponded with two blind Americans, Harvey Lauer and Margaret Butow, who were involved with the development of two new handheld optophones, the visotoner and the stereotoner. Lauer sent a visotoner to Jameson along with recorded instructions on how to use it. In a series of tapes exchanged by mail see Figures 9 and 10 , they weighed the advantages and disadvantages of each. Which code was clearer, the six notes of the British optophone or the nine of the visotoner? Did the portability of the visotoner outweigh the tracking complexities it entailed?

Chapter 2 : 5 Reasons why machine translations will never replace human translators - blog.quintoapp.com

The field of machine translation (MT)â€"the automation of translation between human languagesâ€"has existed for more than fifty years. MT helped to usher in the field of computational linguistics and has influenced methods and applications in knowledge representation, information theory, and mathematical statistics.

Anyone with an interest in the history of computational linguistics will find much to relish and learn from in this weighty collection of articles past. Lest we forget, MT was one of the first nonnumerical applications proposed for the digital computer following the Second World War, and its often tumultuous year history has had a significant impact on the entire field of computational linguistics. Indeed, this very journal can trace its lineage back to the journal whose original title was Mechanical Translation. Though not a proper history of MT, Readings in Machine Translation is certainly a historical collection. For this alone, Nirenburg, Somers, and Wilks deserve our gratitude. The volume begins with the famous memo- randum that Warren Weaver sent out to some professional acquaintances in , which is generally taken to mark the genesis of machine translation; and the most recent paper included dates back to the fourth MT Summit in The editors cite three: Well, as criteria go, that certainly sets a high standard! And yet many of these articles seem to meet it with ease. One reads these papers today, decades after they were written, and one still cannot help but be impressed. Needless to say, not all the articles included in Readings in Machine Translation come up to this high standard; that would be too much to expect. In other cases, one wishes the editors had made more liberal use of their prerogative to abridge. Another reason for the excessive length of Readings in Machine Translation is that the book is divided into three distinct sections, each under the responsibility of one of the editors. There are obvious overlaps between these divisions, in the sense that articles included in one section could just as well fit into another. The editors acknowledge this, and in itself it is not very serious. In his introduction, for example, Nirenburg cites numerous, often lengthy passages from the articles by the early MT pioneers that purportedly support his preferred approach to meaning-based MT. A more serious criticism of Readings in Machine Translation is that the book is some- what dated. This is a rather paradoxical charge for a collection of historical articles; what I mean by it is this: In fact, I was sent a preliminary version by the publisher in In the last few years, for example, there has been an impressive resurgence of activity in machine transla- tion, particularly in the United States, where statistical methods drawn from speech recognition and various techniques borrowed from machine learning have proven re- markably successful. Had the editors been more aware of the profound impact of these new influences on the field, they would perhaps have modified their selection of articles. As it is, only two of the thirty-six papers in the collection explicitly address data-driven or statistical methods in MT: Which brings me to my final criticism of this otherwise wonderful volume. Watson Research Center in the late s that eventually produced the Mark I system, later installed at the U. And where did the article included in this collection first appear? To find the answer to that question, and indeed to locate the source of Book Reviews all thirty-six papers included in Readings in Machine Translation, one has to search the source notes that appear at the end of the volume: It would have been so much easier and more helpful to display this information on the first page of each contribution! Indeed, one wishes the editors had seen fit to include a short introductory note to each article, providing a few words of historical background on the author, or at least his or her affiliation at the time the paper was published. But these are more or less minor quibbles, and they do not significantly detract from the value of this generous volume: Ellis Horwood, Chichester, England.

Chapter 3 : Download [PDF] Readings In Machine Translation Free Online | New Books in Politics

The second section, on theoretical and methodological issues, covers sublanguage and controlled input, the role of humans in machine-aided translation, the impact of certain linguistic approaches, the transfer versus interlingua question, and the representation of meaning and knowledge.

Statistical machine translation Statistical machine translation tries to generate translations using statistical methods based on bilingual text corpora, such as the Canadian Hansard corpus, the English-French record of the Canadian parliament and EUROPARL , the record of the European Parliament. Where such corpora are available, good results can be achieved translating similar texts, but such corpora are still rare for many language pairs. Generally, the more human-translated documents available in a given language, the more likely it is that the translation will be of good quality. With further development, this may allow statistical machine translation to operate off of a monolingual text corpus. In this approach, the corpus that is used is one that contains texts that have already been translated. Given a sentence that is to be translated, sentences from this corpus are selected that contain similar sub-sentential components. Hybrid machine translation Hybrid machine translation HMT leverages the strengths of statistical and rule-based translation methodologies. The approaches differ in a number of ways: Rules post-processed by statistics: Translations are performed using a rules based engine. Statistics guided by rules: Rules are used to pre-process data in an attempt to better guide the statistical engine. Rules are also used to post-process the statistical output to perform functions such as normalization. This approach has a lot more power, flexibility and control when translating. It also provides extensive control over the way in which the content is processed during both pre-translation e. More recently, with the advent of Neural MT, a new version of hybrid machine translation is emerging that combines the benefits of rules, statistical and neural machine translation. The approach allows benefitting from pre- and post-processing in a rule guided workflow as well as benefitting from NMT and SMT. The downside is the inherent complexity which makes the approach suitable only for specific use cases. One of the proponents of this approach for complex use cases is Omniscien Technologies. Neural machine translation A deep learning based approach to MT, neural machine translation has made rapid progress in recent years, and Google has announced its translation services are now using this technology in preference to its previous statistical methods. Tilde is also providing translation solutions based in neural networks. The broken Chinese sentence sounds like "there does not exist an entry" or "have not entered yet" Main articles: Word sense disambiguation and Syntactic disambiguation Word-sense disambiguation concerns finding a suitable translation when a word can have more than one meaning. The problem was first raised in the s by Yehoshua Bar-Hillel. They can be approximately divided into "shallow" approaches and "deep" approaches. Shallow approaches assume no knowledge of the text. They simply apply statistical methods to the words surrounding the ambiguous word. Deep approaches presume a comprehensive knowledge of the word. So far, shallow approaches have been more successful. Why does a translator need a whole workday to translate five pages, and not an hour or two? There are ambiguities one has to resolve. For instance, the author of the source text, an Australian physician, cited the example of an epidemic which was declared during World War II in a "Japanese prisoner of war camp". Was he talking about an American camp with Japanese prisoners or a Japanese camp with American prisoners? The English has two senses. A shallow approach which simply guessed at the sense of the ambiguous English phrase that Piron mentions based, perhaps, on which kind of prisoner-of-war camp is more often mentioned in a given corpus would have a reasonable chance of guessing wrong fairly often. Non-standard speech[edit] One of the major pitfalls of MT is its inability to translate non-standard language with the same accuracy as standard language. Heuristic or statistical based MT takes input from various sources in standard form of a language. Rule-based translation, by nature, does not include common non-standard usages. This causes errors in translation from a vernacular source or into colloquial language. Limitations on translation from casual speech present issues in the use of machine translation in mobile devices. Related to named entity recognition in information extraction. Name entities, in narrow sense, refer to concrete or abstract entities in the real world including people, organizations, companies, places etc. The

initial difficulty that arises in dealing with named entities is simply identifying them in the text. Consider the list of names common in a particular language to illustrate this â€” the most common names are different for each language and also are constantly changing. Another way to deal with named entities is to use transliteration instead of translation, meaning that you find the letters in the target language that most closely correspond to the name in the source language. There have been attempts to incorporate this into machine translation by adding a transliteration step into the translation procedure. However, these attempts still have their problems and have even been cited as worsening the quality of translation. For example, for "Southern California" the first word should be translated directly, while the second word should be transliterated. However, machines would often transliterate both because they treated them as one entity. Words like these are hard for machine translators, even those with a transliteration component, to process. The lack of attention to the issue of named entity translation has been recognized as potentially stemming from a lack of resources to devote to the task in addition to the complexity of creating a good system for named entity translation. One approach to named entity translation has been to transliterate, and not translate, those words. A second is to create a "do-not-translate" list, which has the same end goal â€” transliteration as opposed to translation. A third approach to successful named entity translation is a class-based model. In this method, named entities are replaced with a token to represent the class they belong to. For example, "Ted" and "Erica" would both be replaced with "person" class token. In this way the statistical distribution and use of person names in general can be analyzed instead of looking at the distributions of "Ted" and "Erica" individually. A problem that the class based model solves is that the probability of a given name in a specific language will not affect the assigned probability of a translation. A study by Stanford on improving this area of translation gives the examples that different probabilities will be assigned to "David is going for a walk" and "Ankit is going for a walk" for English as a target language due to the different number of occurrences for each name in the training data. A frustrating outcome of the same study by Stanford and other attempts to improve named recognition translation is that many times, a decrease in the BLEU scores for translation will result from the inclusion of methods for named entity translation. Using these methods, a text that has been translated into 2 or more languages may be utilized in combination to provide a more accurate translation into a third language compared with if just one of those source languages were used alone. If the stored information is of linguistic nature, one can speak of a lexicon. With access to a large knowledge base, systems can be enabled to resolve many especially lexical ambiguities on their own. In the following classic examples, as humans, we are able to interpret the prepositional phrase according to the context because we use our world knowledge, stored in our lexicons: With a large enough ontology as a source of knowledge however, the possible interpretations of ambiguous words in a specific context can be reduced. Other areas of usage for ontologies within NLP include information retrieval , information extraction and text summarization. Because of its size, it had to be created automatically. A definition match algorithm was created to automatically merge the correct meanings of ambiguous words between the two online resources, based on the words that the definitions of those meanings have in common in LDOCE and WordNet. Using a similarity matrix , the algorithm delivered matches between meanings including a confidence factor. This algorithm alone, however, did not match all meanings correctly on its own. A second hierarchy match algorithm was therefore created which uses the taxonomic hierarchies found in WordNet deep hierarchies and partially in LDOCE flat hierarchies. This works by first matching unambiguous meanings, then limiting the search space to only the respective ancestors and descendants of those matched meanings. Thus, the algorithm matched locally unambiguous meanings for instance, while the word seal as such is ambiguous, there is only one meaning of "seal" in the animal subhierarchy. Both algorithms complemented each other and helped constructing a large-scale ontology for the machine translation system. Applications[edit] While no system provides the holy grail of fully automatic high-quality machine translation of unrestricted text, many fully automated systems produce reasonable output. Probably the largest institutional user is the European Commission. In-Q-Tel [49] a venture capital fund, largely funded by the US Intelligence Community, to stimulate new technologies through private sector entrepreneurs brought up companies like Language Weaver. Currently the military community is interested in translation and processing of languages like Arabic , Pashto , and Dari. Machine translation applications have

also been released for most mobile devices, including mobile telephones, pocket PCs, PDAs, etc. Due to their portability, such instruments have come to be designated as mobile translation tools enabling mobile business networking between partners speaking different languages, or facilitating both foreign language learning and unaccompanied traveling to foreign countries without the need of the intermediation of a human translator. Despite being labelled as an unworthy competitor to human translation in by the Automated Language Processing Advisory Committee put together by the United States government, [52] the quality of machine translation has now been improved to such levels that its application in online collaboration and in the medical field are being investigated. The application of this technology in medical settings where human translators are absent is another topic of research, but difficulties arise due to the importance of accurate translations in medical diagnoses. Evaluation of machine translation There are many factors that affect how machine translation systems are evaluated. These factors include the intended use of the translation, the nature of the machine translation software, and the nature of the translation process. Different programs may work well for different purposes. In certain applications, however, e. Even though human evaluation is time-consuming, it is still the most reliable method to compare different systems such as rule-based and statistical systems. It is certainly true that even purely human-generated translations are prone to error. Therefore, to ensure that a machine-generated translation will be useful to a human being and that publishable-quality translation is achieved, such translations must be reviewed and edited by a human. Such research is a necessary prelude to the pre-editing necessary in order to provide input for machine-translation software such that the output will not be meaningless. Both example-based and statistical machine translation rely on a vast array of real example sentences as a base for translation, and when too many or too few sentences are analyzed accuracy is jeopardized. Researchers found that when a program is trained on , sentence pairings, accuracy actually decreases. Ana Nino of the University of Manchester has researched some of the advantages in utilizing machine translation in the classroom. One such pedagogical method is called using "MT as a Bad Model. Nino cites that this teaching tool was implemented in the late s. At the end of various semesters, Dr. Nino was able to obtain survey results from students who had used MT as a Bad Model as well as other models. Overwhelmingly, students felt that they had observed improved comprehension, lexical retrieval, and increased confidence in their target language. Machine translation of sign languages In the early s, options for machine translation between spoken and signed languages were severely limited.

Chapter 4 : Readings in Machine Translation : Harold L. Somers :

The field of machine translation (MT) -- the automation of translation between human languages -- has existed for more than fifty years. MT helped to usher in the field of computational linguistics and has influenced methods and applications in knowledge representation, information theory, and mathematical statistics.

One of the common ways of getting any translation done is to go through a freelance translator or a translation agency. However, for the past years, there has been an emergence machine translators which are becoming popular every day. Machine translation has evolved amazingly over the past few years. Over the past few years, machine translators have become more accurate with the programs learning more about words and their content. The reason being the drastic change in technology with every industry thinking about artificial intelligence and machine learning. This has made machine translation and automatic translation services become some of the driving forces behind the development of this technology. It is evident that companies like Google and Microsoft are serious about automatic translation. See how you can order a Human Translation Machines are replacing people at different workplaces and it is going to get tougher with time. However, it is not definite that human translators will be fully replaced by machine translators. The question here is how efficient are the machine translators and can they replace human translators? Which one is better? The answer here is simple. Machine translators will never replace human translators not even on the aspect of speed let alone accuracy and efficiency. Machines cannot understand culture There is no way you can program a machine to understand culture. Different cultures in the world have different lexical items that are unique to that specific culture. This is one challenge machines have never been able to overcome and it will be extremely difficult for them too. While this is a significant challenge to the machines, native in-country speakers who are well versed in the languages and understand all the idioms and slang the culture has to offer are skilled enough to find suitable equivalents in the target languages. Remember in some cultures; some words mean one thing to them while the same words in another culture mean an entirely different thing. Different people have different accent meaning their pronunciation is different even if the words mean the same thing. Only a human translator can tell the difference because the machine translator will just do the direct word to word translation. These human translators are soaked up in the culture and are able to pick slang and nuances in the text that machines cannot detect. This is a hindrance to machines because they are not advanced to the level of rendering these nuances accurately, but they can only do word to word translations. This puts human translators at a very high level as compared to the machine translators. These words have to be related to the context to help determine their true meanings and only a human translator can do this. It could also mean the tears shed when someone is sad. This now becomes tricky for the machine translator because it cannot relate the word to the context while at the same time, the machine cannot give both the meanings on one text. If you run a document with this kind of words through a machine translator, it will lead to inaccurate translations which disrupt the flow of the text. The document will be confusing and most of the sentences will lack logical meaning. It is difficult to localize machines for different languages New phrases are being developed in any language depending on the dialect. Machine translators cannot be able to pick on these evolutions in the languages as fast as human translators who will be experts on localization and different dialects of their specialist language. This means that the machine translator will have to be continuously programmed and with the sophisticated algorithms these programs are made up of, this will take time and effort to achieve. On the other hand, the human translator is able to pick up on the development in a language quickly than the machine. The million dollar question now is, what if there is no suitable translation for the particular phrase? Machines cannot replicate style and tone Every document written has different style and tone as compared to the other one. A document can have poetic, funny or persuasive style and tone but when it comes to translation, machines miss them. Only a human translator can be able to match and recreate something similar to the style and tone of that document. The machine translator will lose the intended tone and intricate nuances of the original document thus producing something that is flat and soul-less. Specific types of texts like poetry and argumentative essays, in particular, pose a considerable challenge for the

machine translators. They are unable to capture the mood of the text accurately and the result is a flat translation. The original text now becomes inefficient because the person reading the machine translation has something that is not appealing to their mind. A translation cannot be complete without the human touch Machine translators use artificial intelligence and while it is developing day in day out, it can never match human intelligence. Machine translators are becoming more efficient and the quality of the translations produced are becoming more understandable. However, the need for humans to be part of the machine translation equation is always there. After a translation is churned out by the machine, there has to be copyeditors and proofreaders to ensure the piece of translation is grammatically correct and it is also comprehensible. The human translator will put in the finishing touches to the piece localizing and enhancing it for the target audience. If the piece of text has local references and you need it translated for the Chinese, the references can be changed to fit the new context. This is one ability the machines can never beat or be at par with. The complexity of any language is something only humans can fully understand. This is a fact and even the geniuses behind automatic or machine translations tools have conceded to. This means they are developing these devices with the intention to help the human translators but not to force them out of the field. Bottom line is machine translators can never beat the human translators because no machine translator is perfect. Machine translators have come a long way in a short period, but they are still lacking in certain aspects. These are the aspects that make a translation effective for day to day and business use and this is where human translators beat the machines. Translation is most needed if one is not proficient in the source and target language. At Speakt, everything is factored to give you the best translation of the text you provide. If you can barely understand the text in a target or source language, how can you tell if a piece of writing is well translated? Leave it to Speakt and experience the difference. We have professional native translators who fully grasp how the specific language works and will convert all your documents into something publishable. With Speakt, you just have to insert a file or text via simple and user-friendly form and order a translation. Professional native translators will then get started with the translation immediately and you then receive the translated files in your mailbox. It is that quick and easy, hassle-free process that guarantees you of the best results.

Chapter 5 : Readings in Machine Translation | MIT CogNet

A collection of historically significant articles on machine translation, from its beginnings through the early s. The field of machine translation (MT)â€"the automation of translation between human languagesâ€"has existed for more than fifty years.

Weaver, Warren Translation. Donald Booth eds Machine Translation of Languages: Booth Mechanical translation. Computers and Automation 2 4 pages Erwin Reifler The Mechanical determination of meaning. King Stochastic methods of mechanical translation. Mechanical Translation 3 2 , Yngve A Framework for syntactic translation. Mechanical Translation 4 3 , pp. Reprinted in David G. Yehoshua Bar-Hillel The present status of automatic translation of languages. Rhodes A new approach to the mechanical syntactic analysis of Russian. Mechanical Translation 6, Susumu Kuno A preliminary approach to Japanese-English automatic translation. Lamb On the mechanisation of syntactic analysis. Hays Research procedures in machine translation. The In famous Report. MT News International 14 June, , Silvio Ceccato Correlational analysis and mechanical translation. Margaret Masterman Mechanical pidgin translation: Watanabe English-Japanese Machine Translation. Oldenbourg-Munchen, Butterworths, London, p. The place of heuristics in the fulcrum approach to MT, Lingua 21, Martin Kay The proper place of men and machines in language translation. Whitelock Machine translation as an expert task. Jan Landsbergen Montague grammar and machine translation. Nagao Dialogue translation vs. Text translation - interpretation based approach. Proceedings of the 12th International Conference on Computational Linguistics Budapest , pages Christian Boitet Pros and cons of the pivot and transfer approaches to multilingual machine translation. Goodman Treatment of meaning in MT systems. Wilks Where am I coming from: The reversibility of analysis and generation in natural language processing. Elliston Computer-aided translatrion - a business viewpoint. Translating and the Computer. North-Holland Amsterdam , pages System design Introduction H. Zarechnak Three levels of linguistic analysis in machine translation. Journal of the Association for Computing Machinery 6 1 , Vauquois Automatic translation - a survey of different approaches. Statistical Methods in Linguistics Stockholm , pp. Melby Multi-level translation aids in a distributed system. A multilingual system under development, Computational Linguistics 11, Nagao A framework of a mechanical translation between Japanese and English by analogy principle. Artificial and Human Intelligence, Amsterdam: Roossin A statistical approach to machine translation. Computational Linguistics 16, Wilks The Stanford Machine Translation project. Natural Language Processing New York: Algorithmics Press, pages Victor Sadler The textual knowledge bank: Papers presented to the 13th International Conference on Computational Linguistics Helsinki , pages

Chapter 6 : Readings in Machine Translation - PDF Free Download

A more serious criticism of Readings in Machine Translation is that the book is some- what dated. This is a rather paradoxical charge for a collection of historical.

This is a short life span for a science, but in that period remarkable progress has been made, mirroring the advances in the contributing disciplines of computer science and linguistics. We should clarify at this point what this book is not. This is not an introductory textbook on MT. A number of such texts already exist: We have not attempted to provide an extensive bibliography. We list below a number of major books on MT, including the textbooks already mentioned, and a number of collections of articles. Let us turn to the question of what this book is. We assume that readers come to this collection with some knowledge about MT and its history: Having decided to put together this collection, the next task was to decide what should be included. Some decisions were easy: The collection is divided into three sections. The second section contains papers addressing theoretical and methodological issues: The third section concentrates on system design, and as such overlaps slightly with the previous two sections. Early on in the planning of this collection we made a decision not to include descriptions of individual systems: More recently, some systems are covered by a single publication Copeland et al. For this reason, the articles chosen do not date much beyond the early s: A second important criterion was availability: The third criterion was the personal taste of the editors. Not everyone sees the history of MT in the same way, and certainly the reader will question some of the papers we have includedâ€”and some we have omitted. Another project for another editorial team perhaps would be to make a collection of translations of such papers as David Hays did in the early days1 for Russian MT. And perhaps we are not so far from the day when such a collection of translations could be produced with the aid of an MT system, or at least an MAT system. The collection has taken us more time than we expected to put together and bring to publication. We have done our best in this endeavor, as the list of acknowledgments shows. However, we have been unable to identify the copyright owner, or to get a reply from them, in xiii Introduction some cases: As you can see, the papers have all been reset, and we have taken this opportunity to edit some of them lightly, correcting spelling mistakes and other minor inaccuracies. Putting this collection together has been a labor of love, but a huge labor nonetheless, and one which we could not have completed alone. We would like to thank many people who have helped us along the way, but in particular the following have been especially helpful: Wayne State University, Princeton, N. New York, United States, Bonn, West Germany, Colgate University, Hamilton, N. Carnegie Mellon University, Pittsburgh, Pa. University of Texas at Austin, Austin, Tex. Katholieke Universiteit Leuven, Belgium, 5â€”7 July University College, Chester, England, 23â€”25 August Hakone, Japan, 17â€”19 September Munich, West Germany, 16â€”18 August Kobe, Japan, 20â€”22 July Luxembourg, 10â€”13 July Singapore, 13â€”17 September Montreal, Quebec, Canada, 2â€”5 October Cuernavaca, Mexico, 10â€”14 October Paducheva , Exact Methods in Linguistic Research trans. An Introduction to Machine Translation. Center for the Study of Languages and Information. The State of the Art: Indeed, to give just a few examples, the journal Computational Linguistics, so familiar to us today, started its existence as Mechanical Translation, and was later renamed, in turn, Mechanical Translation and Computational Linguistics and The American Journal of Computational Linguistics before assuming its current name. Also, Prolog, a major programming language, was launched with MT in mind. While MT is an application area, it is surprising that it can hardly be considered a direct application of theoretical or descriptive linguistics. This was painfully obvious in the early days of MT. When they serve his purpose, he will consider them. But he will ignore them when an arbitrary treatment of the language material better serves his purpose [. From the s on, MT was, in fact, often used to apply contemporary linguistic theories, but the systems that were directly inspired by a particular linguistic theory were usually seldom comprehensive or broad-coverage. Parallel to the search for the best underlying method for carrying out translation was the policy to use the best and newest advances in computer hardware and software. MT has been widely considered a tangible goal since the late s, with the advent of the digital computer, the concept of stored program and the promise of large storage devices. Certainly, the wartime successes of cryptography in

the early s in the U. The mathematicians and early computer scientists who made spectacular progress in breaking the enemy codes during the war undertook, riding the wave of spectacular successes, to branch into other endeavors and extend their methods, proven on a complex task, to other areas. Importing a technique or a theory that proved successful or promising in one area into another has always been popular. Translation of natural language seemed to be a very natural extension for the methods used in breaking codes. It is no surprise, therefore, that the treatise universally considered as the major impetus for the original interest in MT proceeds intellectually from the metaphor of cryptography: When I look at an article in Russian, I say: I will now proceed to decode. The pendulum would swing once again in the late s, when the renewed emphasis on results and system evaluation in competition would bring back the engineering methods and attitudes familiar from the early days of MT and often quite detached from the knowledge accumulated in linguistics. It is, indeed, remarkable how little impact theoretical linguistics had on the early machine translation. The new discipline borrowed more not only from cryptography but also from philosophy and mathematical logic. In his book Machine Translation: Past, Present, Future and in the page historical survey article in Machine Translation in , he presents a vivid general picture of the events surrounding the early developments in machine translation. The contributions in this section are in approximate chronological order. The story of the memorandum and the events that both led to it and followed it is well presented in Hutchins If ever there was a case of a well-informed, well-positioned and forward-looking enthusiast almost single-handedly creating the initial momentum for a discipline, it is Warren Weaver with respect to MT. Thus, among other recipients of early grants to carry out experiments in non-numerical applications of computing was Andrew Booth of Birkbeck College of the University of London, who concluded, in late , that MT was a prime area for such an endeavor. It is a very interesting document in that the reader should realize that the work described was truly trail-blazing and pioneering. There was no paradigm of MT research in existence yet, and even though Booth does not present his work in a paradigmatic mode, some tacit assumptions about it are interesting to note. Thus, the following observation about the process of translation sets up the overall view of MT as a process of ambiguity resolution. The determination of intended meaning depends not only on the semantic peculiarities of the source language, but on the semantic peculiarities of the target language as well! As already mentioned, our problem is multiple meaning in the light of source-target semantics. As a matter of fact, it cannot be so used, and therefore a literal translation would be completely unintelligible. Of course, most of the present-day MT systems do not attempt to resolve this type of problem dynamically, and typically are only capable of doing this or even considering this as a problem! From among multiple nongrammatical meanings the translation mechanism will extract the intended meaning by determining the nongrammatical meaning in which two or more syntactically correlated source forms coincide. Experiment shows that if an average person guesses the successive words in a completely unknown sentence he has to be told only half of them. In machine translation the procedure has to be generalized from guessing merely the next word. The machine may start anywhere in the sentence and skip around looking for clues. The six types of [analysis] clues are 1. Recognition of coherent word groups, such as idioms and compound nouns. The syntactic function of each word. The selectional relations between words in open classes, that is, nouns, verbs, adjectives, and adverbs. The ability of the translating program to determine antecedents will not only make possible the correct translation of pronouns, but will also materially assist in the translation of nouns and other words that refer to things previously mentioned. All other contextual clues, especially those concerned with an exact knowledge of the subject under discussion. These will undoubtedly remain the last to be mechanized. Finding out how to use these clues to provide correct and accurate translations by machine presents perhaps the most formidable task that language scholars have ever faced. Attempts to learn how to utilize the above-mentioned clues have followed two separate approaches. This approach is used by those who are seeking a short-cut to useful, if not completely adequate, translations. The other approach concentrates on trying to obtain a complete understanding of each portion of the problem so that completely adequate routines can be developed. In view of this, it is remarkable that Bar Hillel, an eminent philosopher of language and mathematical logician, has never written or designed an MT system. In MT, he was a facilitator and an outstanding intellectual critic. His unusual ability to understand the nature of the various problems in MT and the honesty and evenhandedness of

hisâ€”usually very strongly heldâ€”opinions set him apart from the runof-the-mill system designer, too busy building a system to be able fully to evaluate its worth, or amateur critic who often judges MT by an impossible, though popular standard of the best translations performed by teams of professional human translators, editors, domain specialists and proofreaders. On the 95 percent approach: It is probably proper to warn against a certain tendency which has been quite conspicuous in the approach of many MT groups. These groups, realizing that FAHQT [Fully automated, high-quality MT] is not really attainable in the near future so that a less ambitious aim is definitely indicated, had a tendency to compromise in the wrong direction for reasons which, though understandable, must nevertheless be combated and rejected.

Chapter 7 : Machine translation - Wikipedia

The field of machine translation (MT) -- the automation of translation between human languages -- has existed for more than fifty years. MT helped to usher in the field of computational linguistics and has influenced methods and applications in knowledge representation, information theory, and.

Chapter 8 : OPTOPHONIC READING, PROTOTYPING OPTOPHONES - Amodern

READINGS IN MACHINE TRANSLATION Download Readings In Machine Translation ebook PDF or Read Online books in PDF, EPUB, and Mobi Format. Click Download or Read Online button to READINGS IN MACHINE TRANSLATION book pdf for free now.

Chapter 9 : Readings in Machine Translation

In his book Machine Translation: Past, Present, Future () and in the page historical survey article in Machine Translation in , he presents a vivid general picture of the events surrounding the early developments in machine translation.