

Karolina Puchała-Ladzińska

Uniwersytet Rzeszowski

karolinapuchala1@gmail.com

Translator Google wie najlepiej... Czy na pewno?

O rozwoju i jakości tłumaczenia maszynowego

Wstęp

Jak łatwo można zauważyć, świat nieustannie się zmienia, a ogólny postęp dotyczy niemal wszystkich dziedzin ludzkiej działalności, w tym również dziedziny przekładu. To, co jeszcze kilka dekad temu było zadaniem żmudnym, obecnie jest znacznie ułatwione dzięki technologii. Tłumacze nie muszą już polegać jedynie na swojej wiedzy i kilku dostępnych słownikach, ale dziś mają do dyspozycji ogrom narzędzi, dzięki którym mogą być bardziej produktywni. Narzędzia te są w zasięgu nie tylko profesjonalnych tłumaczy, ale właściwie każdego.

Obecnie każdy, kto ma dostęp do Internetu, jest w stanie zrozumieć tekst napisany w obcym języku dzięki dostępowi do forów językowych, licznych słowników on-line oraz, przede wszystkim, translatorów internetowych, które są w stanie tłumaczyć zarówno pomiędzy popularnymi, jak i bardziej „egzotycznymi” językami. Dużą zaletą takich tłumaczy jest to, że swoje zadanie wykonują natychmiast i można z nich korzystać bez dodatkowych opłat – pod tymi dwoma względami maszyny wciąż mają przewagę nad ludźmi. W ten sposób przyczyniają się również do eliminacji jednej z barier istniejących od niepamiętnych czasów – bariery językowej. Biorąc pod uwagę fakt, że tłumaczenia maszynowe są coraz lepszej jakości, być może bariera ta któregoś dnia zniknie zupełnie.

Niniejszy artykuł omawia problem tłumaczenia maszynowego z uwzględnieniem jego zalet i wad. Głównym celem artykułu jest przetestowanie najpopularniejszego translatora on-line – Google Translate. Fragmenty różnych tekstów zostaną „przepuszczone” przez translator, a wyprodukowane tłumaczenia poddane analizie i porównane z tekstami

wyjściowymi. W ten sposób będzie można określić typy tekstów, z tłumaczeniem których program radzi sobie dobrze, a których nie potrafi jeszcze tłumaczyć w stopniu zadowalającym.

Tłumaczenie maszynowe: czy maszyna może imitować ludzki mózg?

Choć postawione powyżej pytanie może wywoływać skojarzenia z popularyzowaną przez filmy science fiction wizją okrutnych maszyn buntujących się przeciw swym twórcom i przejmujących władzę nad światem, trudno jest tak naprawdę dać tutaj jednoznaczną odpowiedź. Nie ulega wątpliwości, że już zainwestowano ogromne sumy pieniędzy w rozwój tłumaczenia maszynowego. Zarówno informatycy, jak i językoznawcy od dłuższego już czasu pracują nad systemem, który byłby w stanie przetłumaczyć każdy rodzaj tekstu. Choć pewien sukces na tym polu został już osiągnięty (rozwój systemów takich, jak SYSTRAN, LOGOS, METAL czy system CANDIDE bazujący na modelach statystycznych – wszystkie one potrafią tłumaczyć pomiędzy konkretnymi parami języków), osiągnięte rezultaty wciąż pozostawiają sporo do życzenia. Jak dotąd nie wiadomo, czy komputer kiedykolwiek będzie w stanie tłumaczyć na poziomie zbliżonym do człowieka.

Koncepcja maszyny imitującej działanie ludzkiego umysłu wydaje się teoretycznie możliwa, biorąc pod uwagę hipotezę silnej sztucznej inteligencji zaproponowanej przez Johna Searle'a. Hipoteza ta zakłada, że ludzki mózg działa na zasadzie algorytmów i algorytmy te można wprowadzić do komputera tak, aby potrafił on wykonać wszystkie zadania intelektualne typowe dla ludzi, w tym te wymagające myślenia twórczego, jak np. pisanie poezji czy komponowanie muzyki [Penrose, 1996: 28-38]. To oznaczałoby, że któregoś dnia maszyna opanuje sztukę tłumaczenia, które również jest procesem twórczym. Tutaj jednak nasuwają się pewne pytania: czy myślenie twórcze można w ogóle zdefiniować za pomocą algorytmów? Nawet jeśli tak, to czy maszyny będą w stanie wykonywać twórcze zadania tak samo dobrze, jak ludzie? Zakładając, że maszyna będzie wystarczająco „twórcza”, czy możliwe jest zebranie wszystkich informacji składających się na wiedzę o świecie i wprowadzenie ich do pamięci systemu? Tłumaczenie wymaga

nie tylko kreatywnego myślenia, ale również szerokiej wiedzy ogólnej z różnych dziedzin oraz świadomości różnic kulturowych. Czy nie jest to aby nazbyt wiele dla maszyny?

Jak dotąd wszystkie te pytania pozostają bez odpowiedzi. Sebastian Kozłowski [2002: 66] sugeruje, że pozornie proste zadanie przetłumaczenia tekstu z jednego języka na drugi jest w rzeczywistości zbyt skomplikowane, aby można było je w pełni zautomatyzować. Jednak, jak można się spodziewać, poziom trudności będzie tu zależał w dużym stopniu od podobieństwa strukturalnego języków i ich bliskości kulturowej – tłumaczenie maszynowe pomiędzy dwoma językami indoeuropejskimi będzie prawdopodobnie bardziej poprawne niż pomiędzy językami odległymi od siebie geograficznie i kulturowo, przykładowo między polskim i japońskim.

Wracając do pytania postawionego w tytule tej części, należy stwierdzić, że pomimo ogromnego rozmiaru swojej bazy danych, komputer nie jest jak dotąd w stanie naśladować ludzkiego umysłu w sposób na tyle zbliżony, aby nie produkować dziwnych i nonsensownych tłumaczeń, jak w bardzo popularnym i często cytowanym przykładzie, gdzie zdanie biblijne w języku angielskim: *The spirit is willing but the flesh is weak* („Duch wprawdzie ochoczy, ale ciało słabe”) zostało przetłumaczone przez maszynę na język rosyjski w taki sposób, że po ponownym przełożeniu go na angielski rezultat brzmiał: *The vodka is strong, but the meat is rotten* („Spirytus jest mocny, ale mięso nieświeże”) [Hutchins, 1995: 17]. To tylko jeden z wielu przykładów błędów, które zdarzają się, gdy maszyna ma przetłumaczyć zdanie silnie osadzone w kontekście. Choć tego typu błędy są niewątpliwie zabawne, to jednak nie mają prawa się zdarzać, jeśli zależy nam na tłumaczeniu wysokiej jakości, szczególnie dla celów biznesowych, ponieważ taki błąd bardzo poważnie osłabia reputację firmy. Ciekawym przykładem jest tutaj pewna chińska kawiarnia, która padła ofiarą zarówno błędu tłumaczenia maszynowego, jak i ludzkiej ignorancji: kawiarnia potrzebowała szyldu z napisem w języku chińskim i angielskim. Chiński tekst, 餐厅, który oznacza „salę jadalną” został przetłumaczony na angielski jako: „Translate Server Error” („Błąd serwera”)¹. Jak można się domyślić, automatyczny translator nie był w stanie przetłumaczyć chińskiego tekstu i wyświetlił

¹ Przykład pochodzi ze strony <http://blog.lingo24.com/machine-translation-blunders>.

jedynie informację o błędzie, ale ponieważ użytkownik programu nie znał angielskiego, założył, że powstały tekst jest tłumaczeniem chińskiego znaku.

Te oraz inne liczne przykłady gaf tłumaczeniowych popełnianych przez maszyny pokazują, że komputer nie jest na razie w stanie wyprodukować efektywnego tłumaczenia, ponieważ nie posiada tego, co można by nazwać „czynnikiem ludzkim”, oraz brak mu ludzkich umiejętności, takich jak wrażliwość na niuanse językowe czy umiejętność domyślenia się znaczenia. Istnieją jednak teksty, które maszyny potrafią przetłumaczyć w zadowalający sposób – teksty te oraz ich tłumaczenia maszynowe będą analizowane w dalszych częściach artykułu.

Google Translate – rewolucja w dziedzinie tłumaczenia maszynowego

Google Translate jest jednym z najpopularniejszych i najczęściej używanych systemów oferujących usługę automatycznego przekładu pomiędzy różnymi językami. Program został stworzony przez firmę Google Inc. i obecnie potrafi dokonać natychmiastowego tłumaczenia z i na 90 języków [„Google Translator”, on-line]. System obsługuje również mniej znane języki, jak np. zulu, maori czy telugu. Serwis cieszy się rosnącą liczbą użytkowników, która obecnie szacowana jest na ponad miliard². Istotną cechą systemu Google Translate jest to, że nie tylko oferuje on natychmiastowe tłumaczenie wprowadzanego ręcznie tekstu, ale potrafi również przetłumaczyć sfotografowany napis – wystarczy jedynie zrobić zdjęcie smartfonem, a Google natychmiast rozpoznaje tekst na obrazku i automatycznie tłumaczy go na język wybrany przez użytkownika.

Ostatnio dodaną innowacją jest opcja tłumaczenia mowy w czasie rzeczywistym, co oznacza, że użytkownik mówi do aplikacji, a pisemne tłumaczenie jego słów natychmiast pojawia się na ekranie [Griffiths, Woollaston, 2015].

System Google Translate działa w oparciu o modele statystyczne – program wybiera istotne informacje z korpusu ekwiwalentnych słów i wyrażań w językach wyjściowym i docelowym. Na przykład gdy

² Dane pochodzą z roku 2013 [zob. Schulz, 2013].

program ma przetłumaczyć zdanie z jednego języka na drugi, wyszukuje on pasujące do siebie wyrażenia w zbiorze danych i na ich postawie decyduje, jak najlepiej skonstruować zdanie docelowe. Ponieważ system bazuje na dostępności danych w obydwu językach, może napotkać problem z budowaniem poprawnych zdań w którymś z mniej znanych języków, dla którego bazy danych nie są tak zasobne. Tłumaczenie będzie również bardziej poprawne, jeśli języki wyjściowy i docelowy będą podobne pod względem struktury. Dlatego istnieje większe prawdopodobieństwo poprawności tłumaczenia pomiędzy językami angielskim i hiszpańskim, niż choćby pomiędzy angielskim a japońskim.

Ale nawet jeśli Google tłumaczy pomiędzy językami podobnymi strukturalnie, wciąż będą pewne aspekty, z którymi nie będzie mógł on sobie poradzić, jak np. dwuznaczności w tekście, które zazwyczaj giną w tłumaczeniu maszynowym.

Wydaje się, że Google Translate z tłumaczeniem pewnych typów tekstów radzi sobie lepiej niż innych. Celem następnej części jest więc przetestowanie tłumacza za pomocą czterech fragmentów różnych tekstów i porównanie wyników.

Pora przetestować Google Translate

Tekst techniczny³

1. Tłumaczenie na język angielski

Tekst wyjściowy

1. Użycie akcesoriów niepolecanych przez producenta może spowodować uszkodzenie urządzenia, pożar lub obrażenia ciała.
2. Nie korzystaj z urządzenia na wolnym powietrzu.
3. Nie wieszaj przewodu na ostrych krawędziach i nie pozwól, aby stykał się z gorącymi powierzchniami.
4. Nie stawiaj urządzenia w pobliżu kuchni elektrycznych i gazowych, palników, piekarników, itp.

³ Obydwa teksty wyjściowe wykorzystane w tej części pochodzą z instrukcji obsługi opiekacza.

5. Podczas opiekania nie dotykaj gorących powierzchni urządzenia. Do otwierania używaj tylko przeznaczonego do tego celu uchwytu.
6. Przed rozpoczęciem czyszczenia najpierw wyjmij wtyczkę z gniazda sieciowego i odczekaj do całkowitego ostygnięcia urządzenia.
7. Przy otwieraniu grilla podczas opiekania należy uważać na wydobywające się gorące opary po bokach płyt grzejnych!
8. Urządzenie powinno być podłączone do gniazdka sieciowego z bolcem uziemiającym.

Tekst docelowy

1. The use of accessory attachments not recommended by the manufacturer may cause damage to the equipment, fire or injury.
2. Do not use the appliance outdoors.
3. Do not hang the cord on sharp edges and do not let it come into contact with hot surfaces.
4. Do not place the unit near ovens and gas burners, ovens, etc.
5. During the cooking, do not touch the hot surface of the device. To open, use only for this purpose handle.
6. Before cleaning, first remove the plug from the outlet and wait for it to cool device.
7. When opening the grill while toasting be careful of the hot vapor coming out the sides of the heating plates!
8. The device should be connected to an outlet with a grounding pin.

Jak łatwo można zauważyć, tłumaczenie jest zupełnie zrozumiałe i w większości poprawne. Formy trybu rozkazującego zostały zachowane w przekładzie, a terminologia techniczna, kolokacje (np. *damage to*) oraz przedimki użyte zostały prawidłowo. Zaskakujący jest fakt, że translator najwyraźniej potrafi posługiwać się synonimami – choć w tekście wyjściowym wielokrotnie powtarza się słowo „urządzenie”, angielskie tłumaczenie nie zawiera tak wielu powtórzeń, a translator wykorzystał cztery inne słowa, których znaczenie jest podobne: *equipment*, *appliance*, *unit*, *device*. Wydaje się więc, że Google Translate „zwraca uwagę” na różnorodność językową, co jest dość zaskakujące.

W powyższym tłumaczeniu można znaleźć tylko niewielkie niedociągnięcia, takie jak powtórzenie słowa *ovens* w punkcie 4 – wydaje się, że

translator przetłumaczył zarówno „kuchnie elektryczne”, jak i „piekarniki” jako *ovens*, choć istnieje różnica pomiędzy tymi dwoma pojęciami, a lepszym ekwiwalentem tego pierwszego byłyby *electric cookers*. Innym niedociągnięciem jest sformułowanie użyte w punkcie 5: *To open, use only for this purpose handle*. W tym wypadku translator po prostu skopiował oryginalny polski szyk zdania, co sprawiło, że angielskie tłumaczenie jest stylistycznie niepoprawne. Jednak nawet pomimo tej niezręczności zdanie wciąż jest komunikatywne, a czytelnik nie powinien mieć problemu z jego zrozumieniem. Translator skopiował oryginalny szyk zdania również w punkcie 6: *and wait for it to cool device*, nie wiedząc, że *it* oraz *device* odnoszą się do tego samego przedmiotu i dlatego jedno z tych wyrażen jest zbędne. W punkcie 7 natomiast brakuje przyimka: *hot vapor coming out (of/on/from) the sides of the heating plates*.

Pomimo tych wszystkich niedoskonałości, Google Translate całkiem dobrze poradził sobie z przetłumaczeniem tekstu wyjściowego i stworzył zrozumiałe i w dużej mierze poprawny tekst techniczny w języku angielskim.

2. Tłumaczenie na język polski

Tekst wyjściowy

1. Before use, thoroughly read the operation manual.
2. Extreme care is required in case of children presence in vicinity of the appliance.
3. Do not use your roaster for any other purpose except of its assignment.
4. Do not immerse the appliance, its cord and plug in water or any other fluids!
5. Do not leave the roaster without supervision in course of its operation.
6. Always pull out the plug from the socket unless the appliance is used or before its cleaning.
7. Do not use the roaster in case of its damage, also if its cord or plug is damaged – in such case the appliance should be handed over to an authorized service shop for repair.

Tekst docelowy

1. Przed użyciem należy dokładnie przeczytać instrukcję obsługi.

2. Szczególną ostrożność jest wymagana w przypadku dzieci obecności w pobliżu urządzenia.
3. Nie należy korzystać z palarni w jakimkolwiek innym celu, z wyjątkiem jego przeniesienia.
4. Nie wolno zanurzać urządzenia, jego przewodu i wtyczki w wodzie lub innych płynów!
5. Nie należy pozostawiać bez nadzoru w palarni trakcie jego eksploatacji.
6. Zawsze wyciągać wtyczkę z gniazdka, chyba że urządzenie jest używane lub przed jego czyszczeniem.
7. Nie używaj palarni w przypadku jego uszkodzenia, także jeżeli jego przewód lub wtyczka jest uszkodzona – w takim przypadku Urządzenie powinno być przekazany do autoryzowanego punktu usługowego, w celu naprawy.

Tutaj rezultat nie jest tak dobry, jak w poprzednim przypadku. Niektóre punkty tekstu wyjściowego zostały przetłumaczone bezbłędnie (punkt 1), przekład innych nie jest już tak poprawny, niemniej jednak wciąż zrozumiały (punkty 2, 4, 6), a jeszcze inne na pewno wywołają konsternację (punkty 3, 5). Głównym problemem było to, że maszyna „zdecydowała” przetłumaczyć słowo *roaster* jako „palarnia”, która byłaby dobrym ekwiwalentem w innych kontekstach, ale na pewno nie tutaj. Również słowo *assignment* zostało błędnie przetłumaczone jako „przeniesienie”, co w tym kontekście nie ma zupełnie sensu.

Kolejna trudność napotkana przez translator w tym przypadku dotyczyła polskich końcówek fleksyjnych, co jest szczególnie widoczne w punktach 2, 4, 7 tekstu docelowego, w których to przymiotniki i rzeczowniki mają niepoprawne końcówki. Innym problemem, podobnie jak w poprzednio analizowanym tłumaczeniu, była tendencja maszyny do kopiowania szyku wyrazów tekstu wyjściowego, na skutek czego w tekście docelowym zobaczyć można nienaturalnie i stylistyczne niepoprawne sformułowania typu: „w przypadku dzieci obecności w pobliżu urządzenia”. Wreszcie, translator miał również widoczne trudności w stosowaniu wielkich i małych liter – z jakiegoś powodu użył małej litery na początku punktu 2, podczas gdy w pozostałych punktach zawsze zaczynał zdanie od litery wielkiej. W punkcie ostatnim natomiast zupełnie niepotrzebnie użył litery wielkiej w wyrazie „Urządzenie”.

Ogólnie rzecz biorąc, tłumaczenie maszynowe tekstu technicznego z języka polskiego na język angielski było o wiele bardziej udane i poprawne niż tłumaczenie podobnego tekstu z języka angielskiego na język

polski. Być może jest to spowodowane tym, iż baza angielskich tekstów technicznych jest większa niż odpowiadających im tekstów polskich.

Tekst medyczny⁴

1. Tłumaczenie na język angielski

Tekst wyjściowy

Należy zapoznać się z treścią ulotki przed zastosowaniem szczepionki.

- Należy zachować tę ulotkę, aby w razie potrzeby móc ją ponownie przeczytać.
- Należy zwrócić się do lekarza lub farmaceuty, gdy potrzebna jest rada lub dodatkowa informacja.
- Szczepionka ta została przepisana ściśle określonej osobie i nie należy jej przekazywać innym.
- Jeśli nasili się którykolwiek z objawów niepożądanych lub wystąpią jakiegokolwiek objawy niepożądane nie wymienione w ulotce, należy powiadomić lekarza lub farmaceutę.

Tekst docelowy

Please read the package leaflet before vaccination.

- Keep this leaflet, you may need to read it again.
- Ask your doctor or pharmacist if you need any further information.
- This medicine has been prescribed for you and you should not pass it on to others.
- If you get any of the side effects or notice any side effects not listed in this leaflet, please tell your doctor or pharmacist.

W tym przypadku tłumaczenie nie tylko jest zrozumiałe i nie rodzi żadnych problemów z interpretacją, ale również jest poprawne językowo. Translator ponownie zaskakuje swoją zdolnością tłumaczenia na angielski. Był nawet w stanie „domyślić się”, że sformułowanie: „szczepionka ta została przepisana ściśle określonej osobie” odnosi się do

⁴ Obydwa teksty wykorzystane w tej części pochodzą z ulotki informacyjnej szczepionki przeciw grypie.

czytelnika ulotki i przetłumaczył je jako: *this medicine has been prescribed for you*. Translator „wiedział” również, że fragment o możliwych działaniach niepożądanych także odnosi się do czytelnika tekstu pomimo tego, że w zdaniu oryginalnym użyta jest forma bezosobowa. Tego rodzaju niemal ludzka zdolność domyślania się znaczenia jest uderzająca, gdy uświadomimy sobie, że chodzi tu o sztuczną inteligencję i ilustruje ona ogromną skalę postępu w dziedzinie tłumaczenia maszynowego.

2. Tłumaczenie na język polski

Tekst wyjściowy

Flu is a disease that can spread rapidly and is caused by different types of strains that can change every year. Therefore, this is why you may need to be vaccinated every year. The greatest risk of catching flu is during the cold months between October and March. If you or your child was not vaccinated in the autumn, it is still sensible to be vaccinated up until the spring since you or your child runs risk of catching flu until then. The doctor will be able to recommend the best time to be vaccinated. The incubation period for flu is a few days, so in case of exposition to flu immediately before or after your vaccination, you could still develop the illness.

Tekst docelowy

Grypa jest chorobą, która może się szybko rozprzestrzeniać i jest spowodowana przez różne rodzaje odkształceń, które zmieniają się w każdym roku. W związku z tym, dlatego być może trzeba się szczepić co roku. Największe ryzyko zachorowania na grypę jest w czasie zimnych miesięcy w okresie od października do marca. Jeśli Ty lub Twoje dziecko nie zostało zaszczepione jesienią, jest nadal sensowne być szczepione aż do wiosny od Ciebie lub Twoje dziecko będzie ryzyko zachorowania na grypę do tego czasu. Lekarz będzie mógł polecić najlepszy czas, aby być zaszczepione. Okres inkubacji grypy wynosi kilka dni, więc w przypadku ekspozycji na grypę bezpośrednio przed lub po szczepieniu, można w dalszym ciągu rozwijać się choroby.

Podobnie jak w przypadku tłumaczenia tekstu technicznego na język polski, również tutaj przekład polski pozostawia wiele do życzenia. Tak jak poprzednio, translator miał wyraźne problemy z doбором

odpowiednich ekwiwalentów – słowo *strains* przetłumaczył jako „odkształcenia”, co zupełnie nie sprawdza się w tym kontekście. Znaczne trudności pojawiły się również z polskimi końcówkami fleksyjnymi, których maszyna najwyraźniej jeszcze nie opanowała. Choć tłumaczenie jest w większości zrozumiałe, brzmi bardzo nienaturalnie i niezręcznie, a szyk wyrazów i kolokacje zostały „skalkowane” z tekstu wyjściowego (np. wyrażenia: „jest nadal sensowne być szczepione”, „najlepszy czas, aby być szczepione”). Tłumacz miał również problem z interpunkcją, gdyż często nie dawał odstępu po kropce zdaniowej.

Interesującym aspektem tego przekładu jest tłumaczenie zaimka *you*. W drugim zdaniu wydaje się, że tłumacz poprawnie „zinterpretował” ten zaimek jako odnoszący się do większej liczby nieokreślonych osób i posłużył się konstrukcją bezosobową, która, poza kilkoma niezręcznymi elementami wprowadzonymi do zdania, brzmi dobrze w kontekście: „trzeba się szczepić co roku”. Jednak już w zdaniach następnych maszyna „zrezygnowała” z konstrukcji bezosobowych i przetłumaczyła zaimek *you* dosłownie, bezpośrednio zwracając się do czytelnika ulotki. Nasuwa się więc pytanie: co sprawia, że komputer zachowuje oryginalny zaimek w pewnych fragmentach tekstu, a opuszcza go w innych? Ponieważ system bazuje na modelach statystycznych, być może aby przetłumaczyć takie fragmenty, wybiera struktury w języku docelowym, które statystycznie są bardziej powszechne w korpusie stanowiącym bazę danych. Niezależnie jednak od powodu, faktem jest, że tłumaczenie maszynowe tekstu medycznego na język polski nie jest tak udane, jak tłumaczenie podobnego tekstu na angielski.

Tekst informacyjny⁵

1. Tłumaczenie na język angielski

Tekst wyjściowy

Kasyna internetowe posiadają szeroki wachlarz gier i miliony zwolenników. Pierwsze kasyno online powstało w 1995 roku, a jego historia jest silnie

⁵ Obydwa teksty wykorzystane w tej części pochodzą ze zbiorów autorki.

związana z historią samego Internetu. Gra odbywa się w wirtualnym świecie, gdzie zawodnicy pozostają anonimowi, a ich działalność na poszczególnych witrynach jest całkowicie poufna. Odpowiednio zaadaptowane oprogramowania służą ochronie prywatności graczy.

Tekst docelowy

Internet casinos have a wide range of games and millions of followers. The first online casino was established in 1995, and its history is closely linked with the history of the Internet itself. The game takes place in a virtual world where players remain anonymous, and their activities at each site is completely confidential. Suitably adapted software are designed to protect the privacy of players.

Jak można zaobserwować w tym przykładzie, podczas tłumaczenia tekstu informacyjnego na język angielski Google Translate potrafił rozpoznać pewne znane kolokacje („szeroki wachlarz”), zastosował poprawne przedimki i w dużej mierze poprawne struktury gramatyczne. Można tu jedynie zidentyfikować dwa błędy: translator potraktował słowo *activities* jako rzeczownik w liczbie pojedynczej, ponieważ użył po nim czasownika *is* zamiast poprawnego *are*, i nie rozpoznał niepoliczalności słowa *software*, traktując je jako rzeczownik w licznie mnogiej i używając po nim *are* zamiast *is*. Poza tymi dwoma przypadkami tłumaczenie jest poprawne i zrozumiałe.

2. Tłumaczenie na język polski

Tekst wyjściowy

The Republic of the Fiji Islands is a nation in the South Pacific known for its rich natural resources, scenery and friendly people. Those looking for white-sand beaches and bright blue waters will not be disappointed. The islands are surrounded by soft coral reefs, which provide a breath-taking underwater spectacle. Visitors can enjoy many one-of-a-kind adventures in a nation less expensive than many in the South Pacific.

Tekst docelowy

Republika Wysp Fidzi jest narodem na południowym Pacyfiku, znanego z bogatych zasobów naturalnych, przyjaznych ludzi i krajobrazy. Ci, którzy szukają białych piaszczystych plaż i błękitnych wód nie będzie rozczarowani. Wyspy otoczone są miękkie raf koralowych, które zapewniają zapierające dech w piersiach podwodny spektakl. Goście mogą korzystać z wielu jedyne w swoim rodzaju przygody w kraju, mniej kosztowne niż wielu na południowym Pacyfiku.

Tym razem translator dobrze zastosował wielkie i małe litery, znalazł polskie ekwiwalenty nazw własnych i właściwie je zapisał zgodnie z polskimi zasadami („Republika Wysp Fidzi”, „Pacyfik”). Największy problem znów stanowiła polska fleksja: podczas gdy niektóre przymiotniki odmienione są poprawnie („bogatych zasobów naturalnych, przyjaznych ludzi), odmiana innych wciąż wciąż pozostawia wiele do życzenia („miękkie raf koralowych”, „wielu jedyne w swoim rodzaju przygody”). Mimo to tekst docelowy jest zrozumiały i do zaakceptowania, pod warunkiem że czytelnik szuka jedynie informacji.

Tekst literacki

1. Tłumaczenie na język angielski

Tekst wyjściowy⁶

Wskazówka wielkiego zegara odmierzała kolejne sekundy. Nieprzerwany strumień sekund obojętnie kapiących cienką strużką, ani wolniej, ani szybciej. Jeszcze godzina, a potem: „Dziękujemy, proszę odłożyć długopisy”. Odda kartkę, na której widnieje jej nazwisko, temat egzaminu, pod spodem tulipan, kilka gwiazdek i dwa samoloty. Tyle nabazgrała przez półtorej godziny, nerwowo rozmyślając, jak wybrnąć.

⁶ Wykorzystany fragment pochodzi z opowiadania Izabeli Sowy *Egzamin* [2004: 137].

Tekst docelowy

Tip of the great clock metered more seconds. A steady stream of seconds matter dripping thin stream, or slower, or faster. Another hour, and then: "Thank you, please put down pens." Odd card, bearing her name on the exam, underneath tulip, a few stars and two airplanes. So much scribbled and a half hours, nervously thinking about how to get out.

Tłumacząc tekst literacki, Google Translate napotkał więcej problemów niż w poprzednio analizowanych typach tekstów. Po pierwsze, mylił słowa – nie rozpoznał, że słowo „wskazówka” użyte na początku pierwszego zdania odnosi się do elementu zegara i pomylił je z *tip*. W efekcie pierwsze przetłumaczone zdanie jest nonsensowne. W zdaniu drugim translator niespodziewanie wprowadził słowo *matter*, czerpiąc je prawdopodobnie z bazy danych ze sformułowania *it doesn't matter* wyrażającego obojętność, które jednak nie może być traktowane jako ekwiwalent polskiego „obojętnie”.

Komputer nie rozpoznał również słowa „odda”, które jest czasownikiem w trzeciej osobie liczby pojedynczej czasu przyszłego i powinno być przetłumaczone jako *she will give back* lub dokładniej – jako że odnosi się ono to karty egzaminacyjnej – *she will hand in*. Również słowo *card* użyte jest niepoprawnie, gdyż w odniesieniu do karty egzaminacyjnej stosuje się słowo *paper*.

W tym tekście translator miał również problem z użyciem zaimka *you* – podczas gdy w przykładzie poprzednio analizowanych typów tekstu komputer stosował go zazwyczaj poprawnie, tutaj czasami zapominał wprowadzić ten zaimek w miejscu, w którym jest on wymagany: *please put down (your) pens*. Translator nie poradził sobie też z metaforycznym opisem w drugim zdaniu tekstu – jego przekład jest bezbarwny i niegramatyczny. Chociaż początek zdania brzmi dobrze, dodatkowe słowo *matter*, które nie powinno tam się znaleźć, powtórzenie słowa *stream* i niepoprawne gramatycznie sformułowanie *or slower or faster* sprawiają, że całe tłumaczenie tego zdania jest nieudane.

2. Tłumaczenie na język polski

Tekst wyjściowy⁷

In fact there are three freely convertible currencies in the Galaxy, but none of them count. The Altairan Dollar has recently collapsed, the Flaninian Pobble Bead is only exchangeable for other Flaninian Pobble Beads, and the Triganic Pu has its own very special problems. Its exchange rate of eight Ningis to one Pu is simple enough, but since a Ningi is a triangular rubber coin six thousand eight hundred miles across each side, no one has ever collected enough to own one Pu. Ningis are not negotiable currency because the Galactibanks refuse to deal in fiddling small change. From this basic premise it is very simple to prove that the Galactibanks are also the product of a deranged imagination.

Tekst docelowy

W rzeczywistości istnieją trzy waluty swobodnie wymienialne w Galaktyce, ale żaden z nich nie liczy. Altairan Dolar niedawno upadł, Flaninian Pobble koralik jest wymienny tylko dla innych Flaninian Pobble Koraliki i Triganic Pu ma swoje własne problemy bardzo specjalne. Jego kurs ośmiu Ningis jednej Pu jest dość proste, ale od Ningi jest trójkątny gumowy monetę sześć tysięcy osiemset mil, na każdej stronie, nikt nie zbiera wystarczy właścicielem jednego Pu. Ningis nie podlegają negocjacjom waluty, ponieważ Galactibanks odmówić czynienia w błahy małą zmianę. Z tego podstawowego założenia, to jest bardzo proste, aby udowodnić, że Galactibanks są także wytworem obłąkanego wyobraźni.

Jak można było się spodziewać, fragment powieści science fiction okazał się dla translatora Google najtrudniejszy, głównie dlatego, że zawiera on nazwy własne będące jednocześnie neologizmami. Jak widać, translator nie rozpoznał ich jako takich i dosłownie przetłumaczył niektóre ich elementy składowe („Flaninian Pobble Koraliki”), a tam gdzie uznał, że to niemożliwe, po prostu zachował nazwę oryginalną (*Triganic Pu, Ningis, Galactibanks*).

⁷ Wykorzystany fragment pochodzi z książki Douglasa Adamsa *The Restaurant at the End of the Universe* [1996: 272].

Poza tą trudnością translator napotkał tu również problemy, które miał podczas tłumaczenia poprzednio analizowanych typów tekstu: interpunkcja (zbyt duża liczba przecinków; nie zawsze po kropce zdaniowej jest odstęp), fleksja języka polskiego („trójkątny gumowy monetę”, „obląkanego wyobraźni”), kolokacje („wymienny dla”), zgodność rzeczownika i czasownika („Galactibanks odmówić”), dobór właściwych ekwiwalentów (słowo *fiddling* przetłumaczone jest jako „błahy”, podczas gdy w tym kontekście chodziło raczej o „denerwujący”/„irytujący”), utarte zwroty (translator nie rozpoznał utartego wyrażenia *small change* i przetłumaczył dosłownie jego komponenty, tworząc sformułowanie „mała zmiana”, zupełnie odbiegające znaczeniowo od zamierzonego). Poza tym, jak łatwo zauważyć, system nie był w stanie wykorzystać techniki zwanej adaptacją i przeliczyć anglosaskie mile na powszechne w Polsce kilometry.

Biorąc to wszystko pod uwagę, nie ulega wątpliwości, że do przekładu tekstów literackich konieczny jest tłumacz człowiek.

Wnioski: „wujek Google” wie najlepiej – czy tłumaczeniu Google można zaufać?

Podsumowując całą przedstawioną analizę, można stwierdzić, że tłumacz Google radzi sobie dobrze z tekstami technicznymi i medycznymi, a więc tekstami o wysokim stopniu sformalizowania, przewidywalnymi strukturami i terminologią. Maszyna na ogół potrafi je skutecznie i poprawnie tłumaczyć zarówno z angielskiego na polski, jak i z polskiego na angielski. W tym drugim kierunku tłumaczenia rezultaty są jeszcze lepsze. Ta tendencja do produkowania lepszych i bardziej poprawnych przekładów na język angielski jest tendencją ogólną i można ją zaobserwować w przypadku wszystkich analizowanych tłumaczeń.

Jeśli chodzi o teksty informacyjne, translator był w stanie wyprodukować tekst docelowy, który przekazywałby ogólny sens oryginału, ale miał wyraźne problemy z takimi aspektami językowymi, jak polska fleksja czy pewne elementy gramatyki angielskiej.

Wreszcie, translator zupełnie nie był w stanie poprawnie przetłumaczyć tekstów literackich. Choć niektóre z przetłumaczonych zdań były do zaakceptowania, a sens był mniej więcej przekazany, większość zdań była jednak nonsensowna i niezrozumiała. To dowodzi, że teksty

literackie wymagają tłumaczy ludzi, którzy nie tylko potrafią rozpoznać i docenić ich wartość estetyczną, ale również zidentyfikować sens nadany im przez autora, który to sens bardzo często przekazywany jest w sposób niebezpośredni. To człowiek jest jak dotąd jedyną istotą, która potrafi „czytać między wierszami”, w razie potrzeby dostosować tekst docelowy do odbiorców lub zrekompensować brak pewnych nieprzetłumaczalnych aspektów za pomocą innych ciekawych zabiegów językowych. Posiada on wiedzę i umiejętność kreatywnego myślenia, co pozwala mu przezwyciężyć wszelkie możliwe trudności tłumaczeniowe, z którymi maszyna nie jest w stanie sobie poradzić.

Dlatego, jak pokazują rezultaty przeprowadzonej analizy, tłumacze mogą być pewni, że ich pozycja nie będzie zagrożona, przynajmniej nie prędko. Jednak mogą oni stopniowo przyzwyczajać się do myśli, że charakter ich pracy ulegnie zmianie. Ponieważ dwie najważniejsze zalety tłumaczenia maszynowego to ich wysoka szybkość i niski koszt, tłumacze będą mogli korzystać z takich systemów jako narzędzi tłumaczeniowych, aby zwiększyć swoją wydajność. To oznacza, że zamiast tłumaczyć tekst od zera, tłumacze będą mogli zlecić to zadanie maszynie, a następnie sprawdzić powstały przekład i nanieść niezbędne poprawki – w ten sposób będą odgrywać bardziej rolę redaktorów niż typowych tłumaczy. Biorąc pod uwagę fakt, że translator Google potrafi tworzyć w dużej mierze poprawne tłumaczenia tekstów technicznych czy medycznych, tłumacz będzie w stanie w ten sposób zaoszczędzić sporo czasu i nie będzie miał dużo pracy przy sprawdzaniu gotowego przekładu. Oczywiście taka współpraca pomiędzy maszyną a człowiekiem nie będzie równie skuteczna w przypadku każdego typu tekstu, ale wydaje się, że przyniesie dobre rezultaty w przypadku tekstów sformalizowanych, o powtarzających się strukturach i przewidywalnej terminologii.

Bibliografia

- „Google Translator”, [on-line] http://en.wikipedia.org/wiki/Google_Translate – 10.01.2015.
- Griffiths, S., Woollaston, V. (2015), „Google Translate app adds conversation mode to translate SPEECH in real time and signs at the click of a button”, [on-line] www.dailymail.co.uk/sciencetech/article-2908490/Google

- Translate-app-adds-conversation-mode-translate-SPEECH-real-time.html – 10.01.2015.
- Hutchins, J. (1995), „«The Whisky Was Invisible» or Persistent Myths of MT”, *MT News International*, 11, s. 17-18.
- Kozłowski, S. (2002), „Co to jest tłumaczenie maszynowe?”, [on-line] http://kf.mish.uw.edu.pl/kog/kog_seb.pdf – 10.01.2015.
- Penrose, R. (1996), *Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*, przeł. P. Amsterdamski, Wydawnictwo Naukowe PWN, Warszawa.
- Piotrowska, M. (2016), „Intuicja tłumacza w kontekście badań przekładoznawczych”, *Między Oryginałem a Przekładem*, 22, 2, s. 57-71.
- Schulz, T. (2013), „Translate This: Google’s Quest to End the Language Barrier”, [on-line] www.spiegel.de/international/europe/google-translate-has-ambitious-goals-for-machine-translation-a-921646.html – 9.01.2015.

STRESZCZENIE

Tematem artykułu jest problem tłumaczenia maszynowego, jego zalet i wad. Głównym celem jest przetestowanie najpopularniejszego translatora on-line – Google Translate. Fragmenty różnych tekstów zostały „przepuszczone” przez translator, a wyprodukowane tłumaczenia poddane analizie i porównane z tekstami wyjściowymi. W ten sposób będzie można określić typy tekstów, z tłumaczeniem których komputer radzi sobie dobrze, a których nie potrafi jeszcze przełożyć w stopniu zadowalającym. Analiza pozwoli również zastanowić się nad rolą tłumacza w przyszłości.

Słowa kluczowe: przekład maszynowy, Google Translate, hipoteza silnej sztucznej inteligencji

SUMMARY

Google Translate knows best... or does it? On the development and quality of machine translation

The article concerns the problem of machine translation, discussing both its advantages and drawbacks. The author’s main purpose is to put the popular online translator, Google Translate, to the test by means of

various text types. Google translations are then analyzed and compared to their source texts. This will allow one to specify with which text types Google can cope considerably well, and which it cannot handle at a satisfactory level. The analysis will also shed some light on the future role of human translators.

Key words: machine translation, Google Translate, hypothesis of strong artificial intelligence