

Maciej Wilamowski

Wydział Nauk Ekonomicznych Uniwersytetu Warszawskiego

Interface człowiek maszyna – święty Graal

Czego oczekujemy od interface'u człowiek maszyna? Odpowiedź na to pytanie wydaje się prosta, a dodatkowo podpowiada w jakim kierunku podążać będzie rozwój narzędzi do komunikacji z komputerem. Szybkość działania i transparentność to dwie najbardziej pożądane cechy. Wyobraź sobie, że przeglądając zdjęcia pomyślałeś, że kilka z nich dobrze by było wysłać do znajomego, ale wcześniej wypadało by zmniejszyć ich rozdzielczość. Pomyśl jak wiele „kliknięć” musisz obecnie wykonać aby wysłać taki email. Teraz wyobraź sobie, że komunikacja z komputerem i wykonanie tej samej czynności, nie wymaga wykonywania żadnej „fizycznej” czynności, a komputer po prostu wykonuje to o czym przed chwilą pomyślałeś, realizując krok po kroku opisane zadanie, a ty jedynie spoglądasz w monitor i mentalnie przytakujesz wykonywanym czynnościom. Brak konieczności świadomego angażowania się w pojedyncze działania (ręczny wybór odpowiednich zdjęć, wybranie adresu email z książki adresowej, etc.) to właśnie transparentność – święty Graal interface'u człowiek-komputer. Jak daleko jesteśmy obecnie od realizacji tych celów i co nas najbardziej ogranicza?

Klawiatura i myszka oraz ekran dotykowy to podstawowe środki komunikacji z komputerem w dniu dzisiejszym. Czy te narzędzia pozwalają na zrealizowanie przedstawionej powyżej wizji? Odpowiedź na to pytanie nie jest oczywista. Zacznijmy od szybkości. Wprawiony programista z łatwością potrafi z wykorzystaniem klawiatury wprowadzać tekst z szybkością ponad 400 znaków na minutę. Zawodnik StarCrafta II podczas gry osiąga ponad 300 akcji na minutę (gdzie jedna akcja to kliknięcie myszką lub polecenie z klawiatury). Przykład gry komputerowej ma o tyle duże znaczenie, że akcje wykonywane w grze nie są wcześniej zaplanowane (na bieżąco okazuje się co trzeba zrobić) w przeciwieństwie do zwykłego pisania na komputerze. Powyższe przykłady sugerują, że przy wykonywaniu dobrze znanych nam czynności obecne narzędzie komunikacji nie stanowią ograniczenia pod względem szybkości ani transparentności. Należy jednak zwrócić uwagę na fakt, że dotyczy to wyłącznie osób, które zajmują się wykonywaniem pewnych czynności zawodowo. Znaczna część użytkowników nie wykonuje tych samych czynności w sposób tak powtarzalny. W przypadku ekranów dotykowych ich obecne wersje nie pozwalają na wprowadzanie treści z taką szybkością jak klawiatury. Co więc kryje przyszłość? Co zrobić aby osiągnąć święty Graal łączności człowieka z komputerem?

Przede wszystkim pracować intensywnie nad interface'em oprogramowania. W najbliższych latach największy postęp w tej dziedzinie osiągnięty zostanie dzięki zmianom wprowadzonym w GUI (graphical user interface) poszczególnych programów. Warto zwrócić uwagę na fakt, iż od początków swojego istnienia, czyli lat 80-tych, GUI większości programów i systemów operacyjnych pozostało praktycznie niezmienione, jeżeli chodzi o sposób działania (okno, menu górne). Dopiero niedawno wraz z pojawieniem się iPhone'a i rozprzestrzenieniem się ekranów dotykowych zaczęto przywiązywać do tej kwestii więcej uwagi i wprowadzać innowacyjne rozwiązania (MS Office 2007, KDE 4, Gnome 3, Unity, etc.). Dynamiczny rozwój interface'ów graficznych w ostatnich latach wyraźnie pokazuje jak wielki postęp może jeszcze nastąpić w tej dziedzinie (dzięki wizualizacji i pół-automatyzacji pewnych działań – zobacz załączniki). Za kilkanaście lat graficzne interface'y będą prawdopodobnie przypominały obecne „kreatory” (wizards), które oferować będą wybór licznych czynności do automatycznego wykonania, prezentując w tle przybliżony rezultat. W przeciwieństwie jednak do dzisiejszych „kreatorów” wszystkie te czynności będą funkcjonować jako dodatkowy krok możliwy do wykonania, a nie obowiązkowy. Ten rodzaj działania można zilustrować w następujący sposób. Po kliknięciu prawym klawiszem na zdjęcie i podświetleniu opcji wyślij do → poczta email, pojawia się dodatkowe menu, które oferuje zmianę rozdzielczości, kadrowanie i retusz zdjęcia, każdą z opcji możemy wybrać (jeśli się na to zdecydujemy). Po dokonaniu np. kadrowania przygotowane zdjęcie automatycznie znajdzie się jako załącznik w nowej wiadomości. Jeżeli jednak nie chcemy korzystać z dodatkowych opcji nie musimy wykonywać żadnego dodatkowego ruchu, co zapewnia zachowanie transparentności.

Drugi krok który należy wykonać, równie ważny, ale trudniejszy do wykonania ze względu na ograniczenia technologiczne, to komunikacja z komputerem za pomocą myśli. Jak dobra (szczegółowa i szybka) komunikacja tą metodą jest wymagana aby stała się wygodniejsza od myszki i komputera? Do poruszania się z łatwością po większości obecnych GUI potrzebujemy prawdopodobnie około 16 różnych komunikatów (0-9, tak/nie i cztery kierunki), następny poziom, swobodne wprowadzanie tekstu, to około 30 dodatkowych komunikatów (przy założeniu możliwości wysyłania różnych komunikatów równolegle). 50 komunikatów, które komputer musi szybko i precyzyjnie rozróżniać to granica, której przekroczenie w połączeniu z dedykowanym GUI i odpowiednie oprogramowaniem do interpretacji powinno pozwolić na efektywniejsze wykorzystywanie komputera niż obecnie. Jak daleko od tego celu jesteśmy? Analizując stan obecny badań prowadzonych w tej dziedzinie wydaje się, że wizja przedstawiona we wstępie tego raportu w perspektywie kilkunastu lat, może małych kilku dziesięciu, może stać się rzeczywistością. Już w 2008 roku naukowcy z Advanced Telecommunications Research (ATR) Computational Neuroscience Laboratories w Kyoto z wykorzystaniem fMRI osiągnęli znaczący sukces w odczytywaniu obrazów w rozdzielczości 10x10 pikseli bezpośrednio z ludzkiego mózgu (zobacz załącznik). Biorąc pod uwagę fakt iż są to początki badań tej metody oraz fakt iż do analizy sygnałów wykorzystana została prosta metoda statystyczna (sparse logistic regression), a nie program, który najpierw „nauczył” by się sygnałów wysyłanych przez daną osobę (np. sieć neuronowa z nauczycielem) to można nieśmiało przypuszczać, że w perspektywie kilkunastu lat granica 50 wyraźnych komunikatów zostanie osiągnięta. Pewnym ograniczeniem jest wielkość maszyn, które potrafią precyzyjnie obrazować ludzki mózg z wykorzystaniem rezonansu magnetycznego. Obecnie są one bardzo duże oraz drogie zarówno w zakupie jak i eksploatacji. Istnieją oczywiście inwazyjne metody, które pozwalają na łatwiejszy dostęp do informacji płynących z mózgu. Jednak takie metody najprawdopodobniej ze względu na ryzyko z jakim się wiążą najprawdopodobniej nie zdobędą popularności.

Podsumowując za kilkanaście lat interface człowiek-maszyna z całą pewnością opierał się będzie o rozbudowane, „inteligentne”, transparentne GUI, które jeżeli technologia na to pozwoli obsługiwane będzie za pomocą myśli.

Załączniki:

<http://www.youtube.com/watch?v=SSGfS6K7pl0> – Gnome 3 Shell – warto zwrócić uwagę na sposób wyszukiwania aplikacji do uruchomienia, rozmieszczanie aplikacji pomiędzy pulpitemi, czy sposoby zmiany wielkości okien.

<http://www.youtube.com/watch?v=2asoMfn1PBQ> – Photoshop save for web. Automatyczna wizualizacja wykonywanych działań. Porównanie 4 opcji, z których trzy to propozycje przygotowane przez program.

<http://www.cns.atr.jp/dni/en/research-projects/reconstruction-visual-images/> – Odczytywanie obrazów z wykorzystaniem fMRI