

Dasher – Character LM

PA154 Jazykové modelování (4)

Pavel Rychlý

pary@fi.muni.cz

March 23, 2021

Dasher

- autoři: David MacKay, David Ward
- Cambridge University; freeware
- podpora pro vysoce efektivní textový vstup za použití jiných prostředků než standardní počítačové klávesnice
- alternativa pro tisíce lidí s různým fyzickým postižením
- zadávání textu na obrazovce pomocí polohovacího zařízení (myš, joystick ...)
- používá pravděpodobnostní prediktivní jazykový model
- stále se ladí (technologie zůstává stejná)

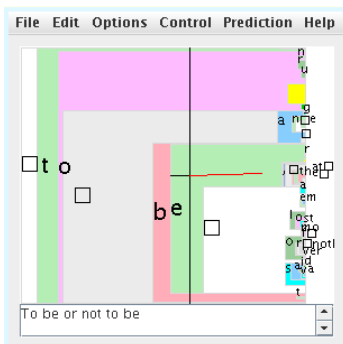
O Dasheru

- Dasher je zdarma
- open-source software
- podléhá požadavkům GNU General Public License
- abeceda pro více než 150 jazyků
- možnost nastavení barvy písma
- systém se učí a nabízí kombinace písmen, které jsou užívanější

Oblasti využití

- asistivní technologie (postižení - bez rukou, s jednou rukou...)
- Pocket PC, iOS, Android, Linux, macOS, Microsoft Windows
- složité jazyky (např. japonština)
- poslední verze 5.0.0 (beta) z 8. dubna 2016

Princip



- písmena v abecedním pořadí, každé písmeno je v obdélníku
- obdélník s vybraným písmenem obsahuje opět úplnou abecedu, ze které lze vybrat 2.symbol atd.
- základní myšlenka: písmena s větší pravděpodobností jsou ve větším obdélníku
- o konkrétním rozložení se rozhoduje na základě jazykového modelu

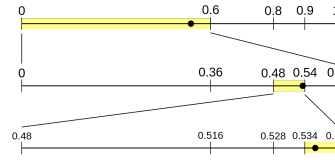
"Inverzní" aritmetické kódování

- aritmetické kódování (kompres textu): kódové slovo je číslo z intervalu (0,1), postupným kódováním symbolů se intervaly zjemňují v poměru pravděpodobnosti výskytu znaku
- metoda bezztrátové komprese dat
- v Dasheru reprezentuje ypsilonová souřadnice celý interval (0,1), kde každý symbol abecedy má přiřazen segment délky odpovídající pravděpodobnosti jeho výskytu v daném kontextu

Aritmetické kódování - příklad pro čtyřsymbolový model

- kódové slovo je číslo z intervalu $[0,1)$
- 60 % pro symbol NEUTRAL; interval je $[0, 0.6)$
- 20 % pro symbol POSITIVE; interval je $[0.6, 0.8)$
- 10 % pro symbol NEGATIVE; interval je $[0.8, 0.9)$
- 10 % pro symbol END-OF-DATA; interval je $[0.9, 1)$
- symbol v sekci END-OF-DATA znamená, že dekódování je kompletní

Aritmetické dekódování

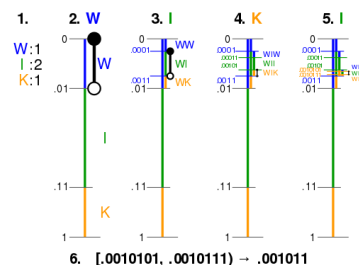


- zpráva je zakódovaná do čísla 0.538
- kódér s intervalem $[0,1)$ se rozdělí na čtyři podintervaly; zpráva je v sekci NEUTRAL
- interval $[0, 0.06)$ se rozdělí na čtyři podintervaly; zpráva je v sekci NEGATIVE
- interval $[0.48, 0.54)$ se rozdělí na čtyři podintervaly; zpráva je v sekci END-OF-DATA

Zakódování zprávy "WIKI" aritmetickým kódováním

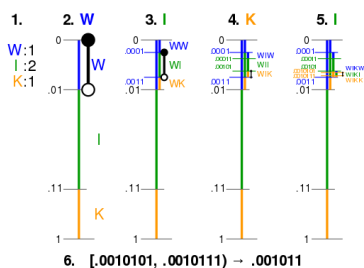
- každý symbol má svoji pravděpodobnost v intervalu $[0, 1)$
- musí být znám počet symbolů zprávy nebo terminální symbol
- interval je znázorněn v binární soustavě

Zakódování zprávy "WIKI" aritmetickým kódováním pokračování



- interval "W" je $[0, 0.01)$
- interval "I" je $[0.01, 0.11)$
- interval "K" je $[0.11, 1)$

Zakódování zprávy "WIKI" aritmetickým kódováním pokračování



- nejdříve kódujeme "W" $[0, 0.1)$
- následně "I" $[0.001, 0.0011)$
- po té "K" je $[0.00101, 0.0011)$
- a nakonec "I" $[0.0010101, 0.0010111)$
- výsledek je číslo z finálního intervalu

PPM (Prediction by Partial Match)

- jazykový model použitý v Dasheru se neomezuje jen na koncept slov
- spojuje informace o n-gramech s pravděpodobnostmi výskytu jednotlivých symbolů ze slovníku
- kontext 4–5 symbolů

PPM - 3 módy

- Standard letter-based PPM (podle částečné shody počítá pravděpodobnost)
- Word-based model (slovník s četností slov)
- Mixture model (PPM/dictionary)

Jazykový model (3)

- jazykový model se učí v čase (učí se nové uživatelské výrazy a obraty)
- vše, co napíšeme, se automaticky ukládá do souboru jako další trénovací data

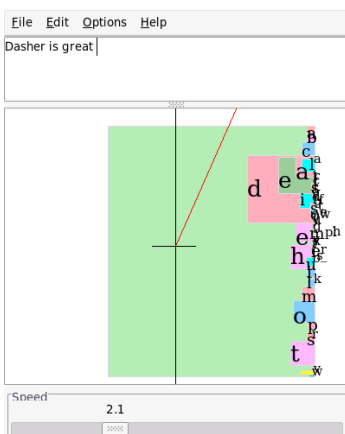
Další vlastnosti

- import trénovacích dat jednoduše načtením souboru
- zdroj dat pro češtinu: Ústav Českého národního korpusu FF UK
- jakékoliv abecedy: např. i LaTeX, C, IPA
- ostatní software – 2 módy: běžné psaní a doplňování slov (uživatel mezi nimi musí přepínat)
- Dasher tyto režimy nerozlišuje

Druhy vstupních metod

- počítačová myš
- touchpad
- touchscreen
- eyetracker
- headmouse
- trackball
- trackpad
- dech
- tlačítka
- snímače náklonu
- ...

Myš, touchpad, touchscreen



- Rychlost vstupu s použitím myši: po 10 minutách tréninku 5–15 slov/min., po hodině 15–25 slov/min., zkušený uživatelé 40 slov za minutu (stejně rychlé jako psaní rukama pomocí klávesnice)
- ukázka Dasheru
- video: ipaq

Eyetracker



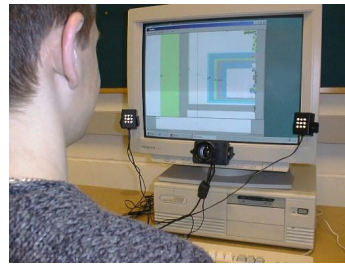
- kamera + senzory snímající, na které místo na obrazovce zaměřuje uživatel pohled
- dřívější cena: 2000 – 4000 USD

Eyetracker



- Tobii Eye Tracker 5
- cena: 229 EUR
- Engineered for gaming
- také vestavěné v (herních) notebookech

Eye Dasher



- rychlost vstupu: po deseti minutách tréninku 7 slov/min., po hodině 20 slov/min., zkušený uživatelé 30 slov za minutu
- eyetracking bez Dasheru, jen s virtuální (on-screen) klávesnicí: 15 slov/min., error-rate 5x vyšší
- poslední položka.

Eye Dasher - Uživatelská přívětivost

- vstup za pomoci virtuální (on-screen) klávesnice je diskrétní (čekání na vypršení časovače, popř. mrknutí)
- Dasher poskytuje spojitý vstup
- video: eye_dasher

Headmouse

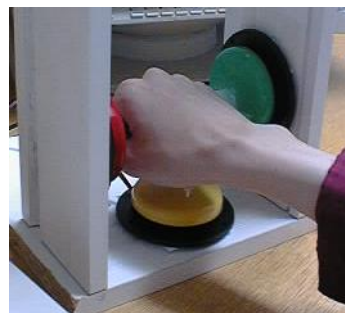
- IR kamera
- reflexivní body
- cena: 500–1500 USD

Breath Dasher



- přímá úměra mezi objemem plic a hodnotou ypsilonové souřadnice
- jednodimenzionální (nelze zpět)
- proto: Control mode
- Control area (Stop, Pause, Move, Delete)
- video: breath_dasher

Button Dasher



- 3 směry
 - dopředu nahoru
 - dopředu dolů
 - zpět

Dasher vs. rozpoznávání řeči

- nepoužitelnost systémů pro automatické rozpoznávání řeči v hlučném prostředí
- i u nejlepších rozpoznávačů asi 5 % chyb (náročná editace chyb)

Speech Dasher: Efficient speech recognition correction

- 1. krok: zadání textu pomocí kombinace řeči a navigace přes ukazovací zařízení (myš)
- 2. krok: řečový rozpoznávač provede počáteční odhad textu, uživatel upraví nebo potvrdí výstup
- počáteční chybovost 22 %, uživatelé většinou vše opraví
- rychlejší než oprava s použitím samostatného rozpoznávání řeči (speciální příkazy)
- rychlejší než samostatný Dasher
- video: [speech_dasher](#)

Další možnosti - Swype

- virtuální klávesnice pro dotykové chytré telefony a tablety
- vyvinuto společností Nuance Communications
- psaní nepřerušovaným tahem na klávesách QWERTY/QWERTZ/AZERTY/National
- odhadování slov pomocí prediktivního slovníku (můžeme doplňovat i vlastní slova)
- větší přesnost pro delší slova (krátká obvykle více možností interpretace tahu na obrazovce)
- zápis bez diakritiky, nabízené varianty s diakritikou

Swype (2)

- rychlost psaní až přes 50 slov/min.
- zvládá i jednoduchou interpunkci (i smajlíky)
- aplikace je schopna učit se z Facebooku, Gmailu, Twitteru...
- dostupný i v češtině
- možnost diktování v různých jazycích pomocí modulu Dragon Dictation (i v češtině)
- video http://www.youtube.com/watch?v=SJ-RAefCG_c

Další možnosti -SwiftKey

- zdarma pro Android, iOS, iPhone
- učí se pomocí předchozí textové komunikace (SMS, Gmail, texty v RSS, přizpůsobuje se i písmenům, která opakovaně mačkáte mírně mimo)
- více jazyků (až 5 zároveň)
- oprava překlepů
- predikce příštího slova (nabízí nejpravděpodobnější varianty následujících slov)
- 800 emoji
- Emoji Prediction feature - učí se predikovat relevantní emoji

SwiftKey (2)

- kvalitní slovníky (odpovídají trendům v komunikaci)
- lze psát ve stylu Swype (tažením) - pod funkcí Flow
- lze zapnout funkce diktování v angličtině
- červen 2012 vydání verze SwiftKey Healthcare; predikce na základě reálných klinických dat
- duben 2016 vydání verze ShakeSpeak; emulování řeči W. Shakespeara k oslavě 400. výročí jeho smrti
- rok 2016 odkoupení SwiftKey společností Microsoft
- video: http://www.youtube.com/watch?v=kA5Horw_SOE

Další možnosti – SlideIT

- podobné klávesnici Swype - psaní tažením mezi znaky
- nižší požadavky na přesnost psaní
- kvalitní slovníky (možnost doinstalovat další vč. češtiny)
- více než 70 jazykových sad
- možnost přizpůsobení klávesnice
- dopočítává varianty slov, které chtěl uživatel napsat
- automatické doplňování mezer a velkých písmen (bohužel chybové)
- video: http://www.youtube.com/watch?v=Tp_7bWuvQwQ

Další možnosti – GO Keyboard

- predikce v mnoha jazycích
- možnost změny skinů a pozadí
- možnost importu jmen a SMS do slovníku
- podpora zadávání textu Swype stylem
- v roce 2017 zjištěn bezpečnostní problém; aplikace zasílala informace o uživateli zpět do Číny (informace o jazyku, lokaci, typu sítě, ...), ovlivněno více než 200 milionů uživatelů
- video: <http://www.youtube.com/watch?v=XQRRvSwpmWc>

Další možnosti

- Perfect keyboard
- TouchPal keyboard
- Google keyboard
- Siine Shortcut keyboard
- Adaptxt keyboard
- ShapeWriter keyboard
- ...