

Penggunaan Pohon Keputusan dalam Pencarian Huruf Kanji

Rizki Alif Salman Alfarisy, 13516005
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
alifalfarisy73@gmail.com, 13516005@std.stei.itb.ac.id

Abstrak – Huruf kanji adalah huruf yang digunakan dalam bahasa Jepang untuk memodelkan suatu konsep dan ide. Untuk dapat menguasai bahasa Jepang, seseorang harus dapat menghafal huruf-huruf kanji ini. Namun, banyak orang yang mengalami kesulitan dalam mempelajari huruf-huruf ini karena jumlahnya yang banyak dan bentuknya yang kompleks. Kamus yang tersedia pun sulit digunakan karena sulitnya menyamakan bentuk huruf dengan daftar yang tersedia di kamus. Makalah ini membahas tentang cara mempermudah pencarian huruf kanji dengan memanfaatkan pohon keputusan. Pohon keputusan menggunakan perbandingan komponen-komponen dari suatu kanji, yaitu jumlah garis dan posisi radikal kanji.

Kata kunci—Kanji, Pencarian, Perbandingan, Pohon.

I. PENDAHULUAN

Bahasa Jepang dan bahasa Mandarin adalah dua bahasa yang cukup terkenal di dunia. Hal ini disebabkan karena kedua negara tempat bahasa ini berasal merupakan negara maju yang sangat berpengaruh di dunia di bidang ekonomi. Produk-produk yang berasal dari Jepang dan Cina tentunya akan melampirkan bahasa dari negara asalnya. Selain itu, kedua negara ini memiliki sifat nasionalisme yang sangat kuat dan tentunya selalu berusaha untuk menyebarkan budaya mereka, salah satunya menggunakan bahasa mereka. Kedua bahasa ini pun mulai dianggap sebagai salah satu bahasa internasional yang penting untuk dipelajari.

Melihat pentingnya kedua bahasa ini, tentulah banyak orang yang ingin mempelajari bahasa ini untuk berbagai tujuan. Namun, bahasa Jepang dan bahasa Mandarin tidaklah mudah untuk dipelajari. Selain karena tata bahasanya yang sangat berbeda dengan bahasa pada umumnya, bahasa-bahasa ini sulit dipelajari karena aksaranya yang sangat banyak. Pada bahasa Mandarin terdapat huruf Han Zi yang jumlahnya puluhan ribu, dan untuk bisa menguasai bahasa Mandarin seseorang minimal harus menghafal sebanyak 6000 huruf. Sedangkan untuk bahasa Jepang terdapat huruf Kanji yang diadopsi dari Cina. Jumlah huruf kanji yang biasa dipakai dalam percakapan sehari-hari di Jepang memang lebih sedikit dibandingkan dengan huruf Han Zi, yaitu hanya 1945 huruf. Selain jumlahnya yang banyak itu, yang membuat huruf-huruf ini sulit dipelajari adalah cara membacanya yang beragam dan kompleksitas bentuknya, bahkan terdapat huruf yang untuk menulisnya dibutuhkan lebih

dari 20 garis.

Tentunya ada banyak usaha untuk mempermudah mempelajari huruf-huruf tersebut salah satunya adalah dengan cara mengumpulkannya ke dalam kamus. Tetapi, walau sudah dikumpulkan menjadi satu buku, mencari huruf yang tepat di buku tersebut sangatlah sulit karena kita harus mencari satu huruf yang sama persis dengan huruf yang kita ingin cari dari ribuan kemungkinan huruf.



Gambar 1. Huruf Kanji untuk kata ‘Kemenangan’

(sumber : <https://it.wiktionary.org/wiki/勝>)

Pada makalah ini, penulis ingin memaparkan suatu cara untuk mempermudah pencarian suatu huruf dengan menggunakan pohon keputusan. Huruf yang akan dibahas dalam makalah ini adalah huruf kanji dari bahasa Jepang karena jumlahnya yang tidak terlalu banyak dan juga tidak sekompleks huruf-huruf Han Zi. Selain itu, huruf kanji biasanya memiliki pola-pola tertentu dalam bentuknya sehingga cenderung lebih mudah untuk dikenali. Pola-pola tersebut akan dimanfaatkan untuk mempercepat pencarian kanji dengan bantuan pohon keputusan.

II. HURUF KANJI

A. Sejarah

Kanji adalah salah satu dari empat set aksara yang digunakan dalam tulisan modern Jepang selain katakana, hiragana, dan romaji. Huruf kanji dipakai untuk melambangkan konsep atau ide (kata benda, kata kerja, kata sifat, atau kata keterangan). Hampir seluruh kosakata selain kata serapan dalam bahasa Jepang bisa dilambangkan dengan Kanji.

Huruf-huruf kanji sebenarnya diadopsi dari huruf Han Zi asal tionghoa yang masuk ke Jepang melalui barang-barang impor pada abad ke-5 Masehi. Sejak saat itu, huruf-huruf Han Zi tersebut sering dipakai pula untuk menulis di prasasti dan

barang-barang di Jepang.

Pada awalnya, orang Jepang hanya meniru huruf-huruf tersebut dan melafalkannya sesuai dengan cara baca bahasa Tionghoa. Tapi lama kelamaan diciptakan sistem *kanbun* (漢文) yaitu cara penulisan huruf Han Zi yang dilengkapi dengan tanda diakritik. Tanda diakritik tersebut ditambahkan untuk menyesuaikan huruf Han Zi agar sesuai dengan tata bahasa Jepang. Setelah itu, huruf-huruf tersebut makin banyak yang dimodifikasi dan disesuaikan cara bacanya sehingga menghasilkan kanji modern yang bentuknya masih mengadopsi bentuk huruf Han Zi tetapi memiliki cara baca yang baru.

B. Komponen Kanji

Setiap huruf kanji memiliki komponen-komponen yang membedakan kanji tersebut dari kanji lainnya. Berikut komponen-komponen penting dari suatu kanji :

1. Cara Pengucapan

Satu huruf kanji dapat memiliki cara baca yang beragam. Setiap kanji memiliki dua cara pengucapan, yaitu *On'yomi* dan *Kun'yomi*.

On'yomi adalah cara mengucapkan huruf Kanji yang mengikuti cara baca asli dari tionghoa saat huruf tersebut pertama kali dikenalkan di Jepang. Cara pengucapan kanji menurut bahasa Tionghoa bergantung pada era dimana kanji tersebut dikenalkan sehingga sebagian besar huruf Kanji memiliki banyak *On'yomi*.

Kun'yomi adalah cara mengucapkan huruf Kanji yang berasal dari Jepang. Pada cara pengucapan ini, setiap huruf kanji hanya memiliki satu cara pengucapan sebagai bentuk usaha untuk membakukan cara membaca kanji.

2. Garis

Huruf-huruf kanji dibentuk dengan cara menorehkan garis-garis dengan aturan tertentu. Garis-garis yang membentuk suatu kanji dapat berupa garis lurus ataupun garis melengkung atau bahkan garis pendek yang menyerupai titik. Terdapat aturan untuk menulis garis-garis tersebut, seperti arah tarikan garis dan bentuk lengkungan. Oleh karena itu, garis-garis pada kanji yang berbeda mungkin memiliki kemiripan. Selain itu, tiap kanji juga mungkin memiliki jumlah garis yang berbeda-beda.

3. Radikal

Radikal kanji adalah bagian dari suatu kanji yang sering muncul di banyak kanji. Radikal dari suatu kanji biasanya memiliki arti yang berhubungan dengan kanji itu sendiri. Misalnya kanji 汁 [sup], 汚 [kotor], 泡 [gelembung], dan 泳 [berenang] sama-sama memiliki 2 titik dan 1 garis di bagian kirinya yang merupakan radikal kanji untuk air.



Gambar 2. Kanji dengan radikal air

(sumber : <http://www.sljfaq.org/afaq/radicals.html>)

Terdapat 214 radikal kanji yang diakui di Jepang. Radikal-radikal inilah yang digunakan untuk mempermudah mengenali suatu kanji. Radikal dari suatu kanji biasanya terletak di bagian tertentu dari karakter kanji, yaitu di bagian kiri [Hen], kanan [Tsukuri], atas [Kanmuri], bawah [Ashi], kiri-atas [Tare], kiri bawah [Nyō], atau mengelilingi karakter [kamae]. Tiap kanji dapat memiliki lebih dari satu yang merupakan kombinasi dari posisi-posisi yang sudah disebutkan sebelumnya.

Hen		私 泳 紙
Tsukuri		利 歌 順
Kanmuri		安 等 雪
Ashi		児 思 炭
Tare		病 庭 厨
Nyō		道 延 麵
Kamae		回 間 式

Gambar 3. Posisi-posisi Radikal Kanji

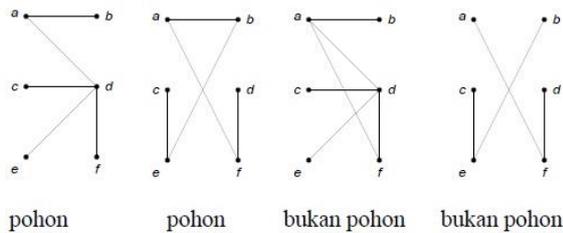
(sumber : <http://www.sljfaq.org/afaq/radicals.html>)

III. POHON

A. Definisi

Pohon adalah graf yang khusus. Misalkan $G = (V, E)$ adalah graf tak berarah-sederhana dan jumlah simpulnya n . Maka, syarat agar G merupakan suatu pohon adalah :

1. Setiap pasang simpul di dalam G terhubung dengan lintasan tunggal.
2. G terhubung dan memiliki $n-1$ buah sisi.
3. G tidak mengandung sirkuit.
4. Penambahan satu sisi pada graf hanya akan membuat satu sirkuit
5. Semua sisi graf adalah jembatan (jembatan adalah sisi yang bila dihapus menyebabkan graf terpecah menjadi 2 komponen).

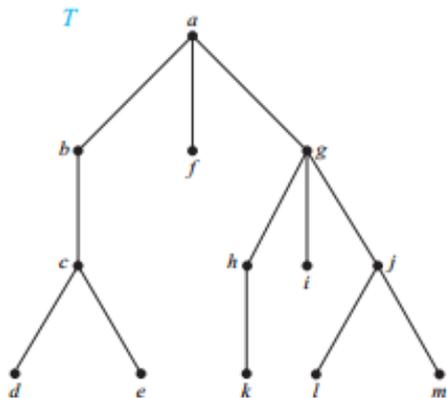


Gambar 4. Contoh Pohon dan bukan pohon
(sumber : Munir, R. *Matematika Diskrit*. Bandung: Informatika Bandung, 2005.)

B. Terminologi pada pohon.

Pohon yang sebuah simpulnya diperlakukan sebagai akar dan sisi lainnya diberi arah sehingga menjadi graf berarah disebut pohon berakar. Berikut ini terminologi penting yang terdapat pada pohon berakar.

- Akar**
Akar adalah sebuah simpul yang memiliki derajat masuk nol. Akar adalah permulaan dari suatu pohon dan hanya memiliki sisi yang keluar dari simpul.
- Anak (child), Orang tua (parent), Saudara (sibling)**
Misalkan T adalah pohon berakar, dan v adalah simpul pada T selain akar. Maka, orang tua dari v adalah simpul unik u dimana terdapat sisi dari u ke v. Apabila u adalah orang tua dari v maka v adalah anak dari u. Simpul yang memiliki orang tua yang sama disebut saudara.



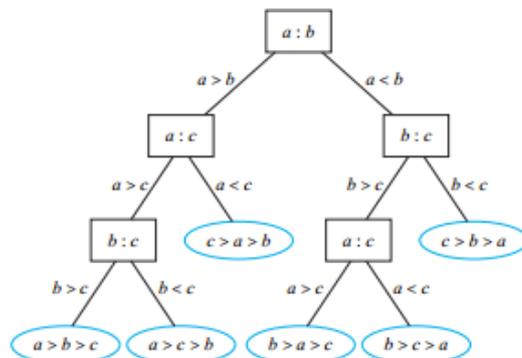
Gambar 5. Pohon berakar
(sumber : *Discrete Mathematics and its application*, Kenneth H. Rosen : 753)

- Leluhur (ancestor) dan Keturunan (descendant)**
Leluhur dari suatu simpul selain akar adalah semua simpul yang dilalui lintasan dari akar ke simpul tersebut, termasuk di dalamnya akar. Sedangkan keturunan dari suatu simpul v adalah semua simpul yang memiliki v sebagai leluhurnya.
- Daun (leaf) dan Simpul Internal (Internal Vertices)**
Daun adalah suatu simpul dalam pohon berakar yang tidak memiliki anak. Sedangkan simpul internal adalah simpul yang memiliki anak.

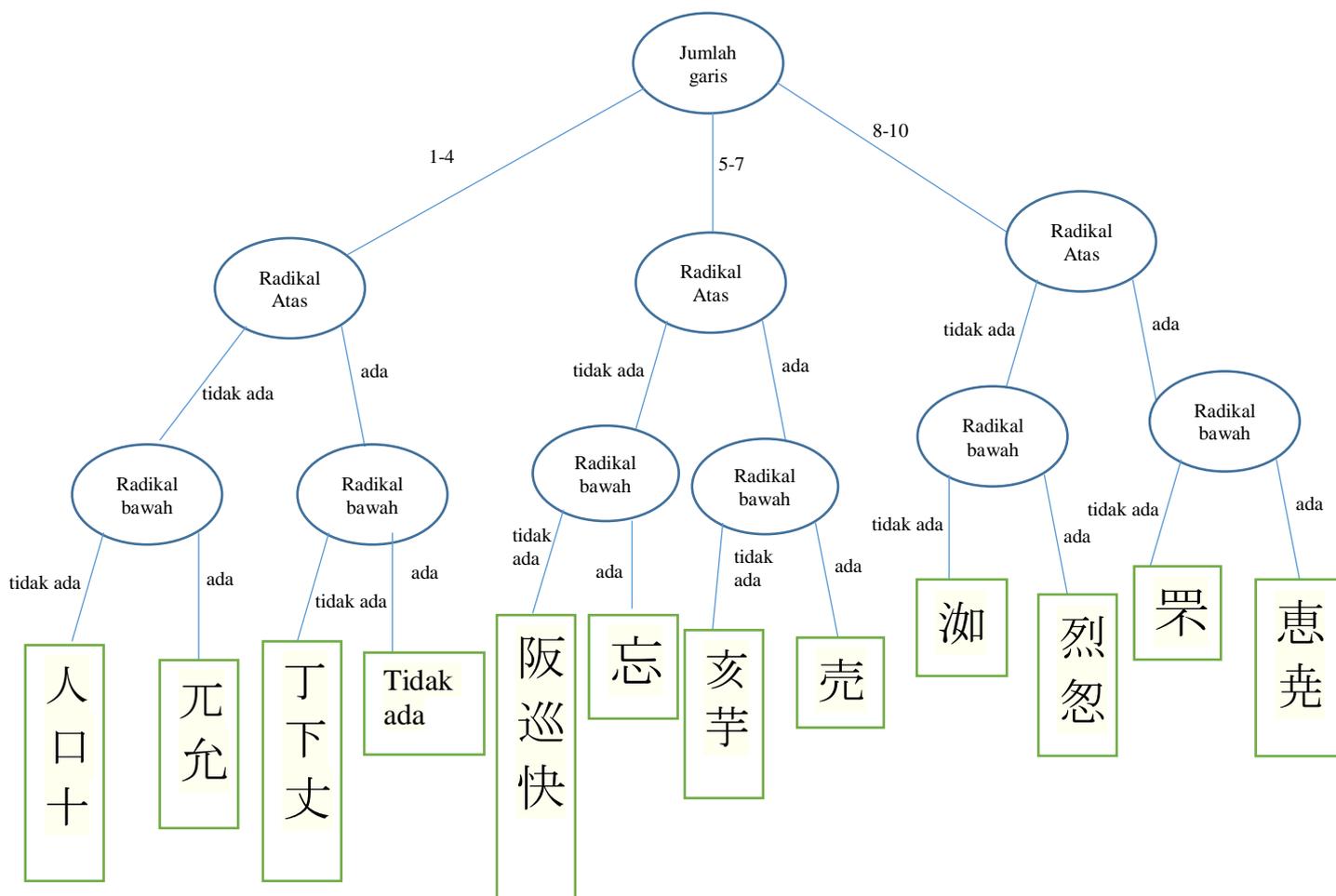
- Upapohon (subtree)**
Jika v adalah suatu simpul dalam pohon berakar T, Upapohon dengan a sebagai akarnya adalah upagraf dari pohon T yang mengandung a dan seluruh simpul keturunannya beserta dengan sisi-sisi yang menghubungkannya.
- Derajat (degree)**
Derajat sebuah simpul pada pohon berakar adalah jumlah upapohon yang dimiliki oleh simpul tersebut.
- Aras (level)**
Akar mempunyai aras nol, sedangkan aras simpul lainnya adalah panjang lintasan dari akar ke simpul tersebut ditambah dengan satu.
- Tinggi (Height)**
Tinggi dari suatu pohon berakar adalah aras maksimum dari pohon tersebut.
- Pohon-n ary**
Pohon berakar disebut dengan pohon n-ary apabila setiap simpul dalamnya memiliki anak tidak lebih dari n. Suatu pohon disebut pohon n-ary penuh apabila setiap simpul dalamnya memiliki tepat n anak. Pohon n-ary dengan n=2 biasa disebut pohon binari.
- Pohon Terurut**
Pohon terurut adalah pohon yang anak dari simpul dalamnya terurut dari kiri ke kanan.

C. Pohon Keputusan

Pohon keputusan (*Decision Tree*) adalah salah satu bentuk aplikasi dari pohon berakar. Pohon berakar yang setiap simpul internalnya berkorespondensi dengan suatu keputusan dan anak-anaknya adalah kemungkinan dari hasil keputusan disebut pohon keputusan. Seluruh kemungkinan solusi dari suatu masalah dimodelkan dengan lintasan menuju ke daun dari pohon ini.



Gambar 6. Pohon keputusan untuk mengurutkan 3 elemen yang berbeda
(sumber : *Discrete Mathematics and its application*, Kenneth H. Rosen : 761)



Gambar 7. Pohon keputusan sederhana untuk pencarian kanji

IV. PENGGUNAAN POHON KEPUTUSAN DALAM PENCARIAN HURUF KANJI

A. Pohon keputusan pencarian huruf Kanji

Untuk mencari huruf kanji yang kita inginkan, digunakan komponen-komponen kanji yang terlihat secara visual untuk mencari kanji. Komponen tersebut adalah jumlah garis dan radikal-radikal kanji.

Pada simpul di Aras 0 atau akar, komponen yang pertama dibandingkan adalah jumlah garis pada huruf kanji tersebut, karena jumlah garis dari suatu kanji adalah komponen yang paling mudah untuk dicek. Selanjutnya sisi-sisi yang menghubungkan akar ini dengan anak-anaknya adalah hasil dari perbandingan, yaitu jumlah garis itu sendiri. Sisi ini bisa merepresentasikan jumlah tepat dari garis atau bisa berupa *range* (1-5,6-10,11-15, dst). Penggunaan representasi *range* dapat kita pilih apabila kita tidak ingin jumlah anak yang terlalu banyak.

Setelah perbandingan jumlah garis pada suatu kanji, pada simpul di aras-aras berikutnya, dilakukan perbandingan dengan radikal-radikal kanji untuk setiap kemungkinan posisi radikal kanji. Misal, pada aras satu dilakukan perbandingan untuk

radikal kanji bagian kiri, maka sisi-sisi yang menghubungkan simpul pada aras satu ke simpul-simpul pada aras berikutnya adalah representasi dari hasil perbandingan radikal kanji bagian kiri. Hasil perbandingan dapat berupa semua kemungkinan radikal kanji bagian kiri atau sekedar ada atau tidak ada. Jika hasil perbandingan adalah semua kemungkinan radikal kanji maka akan menyebabkan pohon semakin lebar tapi hasilnya akan semakin akurat, sedangkan apabila hasilnya hanya ada dan tidak ada maka pohon bisa diperkecil ukurannya namun hasilnya menjadi kurang akurat. Pemilihan representasi ini dapat disesuaikan sesuai dengan kemampuan dan kebutuhan pengguna.

Untuk simpul pada aras-aras berikutnya dilakukan perbandingan radikal kanji untuk setiap posisi berikutnya dengan sisi-sisinya yang merupakan hasil representasi perbandingan seperti yang sudah dicontohkan pada kasus aras satu. Karena terdapat 7 posisi radikal kanji, maka pohon yang terbentuk adalah pohon dengan tinggi 7, yang terdiri 1 aras untuk membandingkan jumlah garis pada Kanji dan 7 aras untuk membandingkan radikal kanji di tiap posisi.

B. Analisis pohon keputusan pencarian kanji

Pada gambar 7, terdapat pohon keputusan sederhana yang digunakan untuk mencari kanji yang memiliki jumlah garis 1-10

kata dan dilakukan perbandingan dengan radikal di posisi atas dan bawah sehingga pohon tersebut memiliki tinggi 3 dan mempunyai 12 daun. Pohon tersebut adalah pohon sederhana dari pohon pencarian kanji.

Misalkan kita ingin mencari kanji tahu arti dari huruf 壳, pertama-tama hitunglah terlebih dahulu jumlah garis pada huruf tersebut, yaitu 7. Karena 8 berada pada range 5-7 maka dari akar, kita telusuri sisi yang berada di tengah, lalu di simpul tersebut akan dibandingkan radikal bagian atas dari kanji tersebut. Kanji tersebut ternyata memiliki radikal bagian atas yaitu 十, Oleh karena itu kita lanjut telusuri sisi ada yaitu sisi kanan dari simpul tersebut. Berikutnya, kita sampai di simpul radikal bagian bawah, kita perhatikan kembali kanji yang kita sedang cari, kanji tersebut ternyata memiliki radikal kanji bagian bawah, yaitu 儿. Karena itu kita lanjut mengikuti sisi bagian kanan dan akhirnya sampai ke solusi yang kita inginkan.

Selain digunakan untuk mengidentifikasi kanji yang belum kita ketahui info selain bentuknya, pohon tersebut juga dapat digunakan untuk mencari kanji dengan komponen yang kita inginkan. Misal, pengguna pohon menginginkan untuk mencari kanji yang jumlah garisnya 1-4 buah, memiliki radikal atas, tetapi tidak memiliki radikal bawah. Pengguna tinggal menelusuri pohon sesuai komponen yang telah ditentukan dan akan menemukan kemungkinan kanji-kanji yang memenuhi kondisi yang ditentukan oleh pengguna.

Pohon keputusan pada gambar 7 bukanlah pohon pencarian kanji ideal karena pada daun terdapat banyak kemungkinan solusi. Hal ini disebabkan karena jumlah sisi yang digunakan pada setiap simpul hanya dimunculkan secukupnya untuk memberikan gambaran mengenai alur penggunaan pohon. Pohon yang sebenarnya berukuran jauh lebih besar dari pohon tersebut.

C. Kelebihan dan kelemahan penggunaan pohon keputusan untuk pencarian kanji

Kelebihan dari pohon ini adalah dapat mempercepat pencarian kanji dan hanya membutuhkan sedikit informasi, yaitu bentuk huruf kanji itu sendiri. Pengguna pohon hanya membutuhkan komponen visual dari kanji yang dicari tanpa perlu mengetahui cara baca ataupun makna dari kanji tersebut. Selain itu, dengan menggunakan pohon pencarian kanji ini, pengguna juga dapat meningkatkan kemampuan menghafal kanji mereka dengan mengingat-ingat radikal kanji di setiap sisi yang pengguna lewat saat menelusuri pohon keputusan ini.

Kelemahan pohon pencarian kanji ini adalah untuk mencapai akurasi yang maksimum, dibutuhkan pohon yang sangat lebar. Pohon tersebut akan sangat sulit untuk dibuat oleh manusia dan tentunya untuk menelusurinya akan memakan waktu lama. Solusi yang penulis tawarkan untuk mengatasi kelemahan ini adalah dengan menggunakan range pada perbandingan jumlah garis dan menggunakan boolean 'ada' dan 'tidak ada' pada perbandingan radikal. Namun, dengan melakukan hal tersebut maka akurasi dari pencarian kanji akan menurun dan solusi dari

pencarian akan menghasilkan banyak kemungkinan seperti pada gambar 7.

Solusi lain untuk kelemahan ini adalah menggunakan program untuk membuat dan menelusuri pohon. Namun yang harus diperhatikan adalah bagaimana cara program melakukan perbandingan. Program harus dapat melakukan pemrosesan citra untuk menghitung jumlah garis pada kanji dan membedakan bentuk-bentuk radikal atau user harus memasukkan secara manual data-data yang dibutuhkan untuk menelusuri pohon keputusan.

V. KESIMPULAN

Pohon sebagai bentuk graf khusus memiliki banyak kegunaan, salah satunya adalah pohon keputusan yang dapat memodelkan persoalan dan memberikan solusi berdasarkan perbandingan-perbandingan.

Salah satu penggunaan dari pohon keputusan adalah untuk melakukan pencarian suatu hal dengan menggunakan deskripsi terkait barang tersebut. Salah satu contohnya adalah dalam dunia literatur dan pendidikan, pohon keputusan dapat digunakan untuk mempermudah pencarian huruf dengan kompleksitas yang cukup rumit dan jumlah yang sangat banyak seperti huruf Kanji dari bahasa Jepang.

Penggunaan pohon keputusan pencarian kanji ini dapat dikembangkan lebih jauh jika dapat dilakukan penelusuran otomatis menggunakan suatu program. Pencarian juga dapat dibuat makin praktis apabila program tersebut dapat melakukan pemrosesan citra untuk mengidentifikasi komponen-komponen kanji yang dibutuhkan untuk melakukan perbandingan pada pohon keputusan.

VI. UCAPAN TERIMA KASIH

Pertama-tama, Penulis mengucapkan syukur kepada Tuhan Yang Maha Esa atas rahmatnya sehingga Penulis dapat menyelesaikan makalah ini. Penulis juga berterima kasih kepada orang tua dan teman-teman Penulis yang telah memberikan dukungan selama penulisan makalah ini baik secara langsung ataupun tidak langsung. Penulis juga mengucapkan terima kasih kepada Ibu Harlili, selaku dosen dari mata kuliah IF2120 Matematika Diskrit yang mengajarkan saya ilmu-ilmu yang saya gunakan dalam makalah ini. Terakhir, Penulis meminta maaf atas kesalahan-kesalahan yang mungkin tidak disengaja dalam makalah ini. Semoga makalah ini dapat bermanfaat bagi orang yang membacanya.

REFERENCES

- [1] R. Munir, Matematika Diskrit, 3rd ed. Bandung: Penerbit INFORMATIKA Bandung, 2010.
- [2] Rosen, Kenneth H., Discrete Mathematics and Its Applications, New York: McGraw-Hill International, 2012, 7th ed
- [3] <https://kanjialive.com/214-traditional-kanji-radicals/> diakses pada 3 Desember 2017 pukul 19.47

- [4] <http://www.sljfaq.org/afaq/radicals.html> diakses pada 3 Desember 2017 pukul 17.43
- [5] <https://it.wiktionary.org/wiki/%E5%8B%9D> diakses pada 3 Desember 2017 pukul 17.50
- [6] https://www.kompasiana.com/kresnap/kanji-huruf-jepang-yang-ruwet-nilimet_550dcf75a333116e1c2e3d88 diakses pada 3 Desember 2017 pukul 17.00

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 3 Desember 2017



Rizki Alif Salman Alfarisy / 13516005