

Yetkin Bul ve Değiştir

Figen Şentürk
İTÜ Bilgisayar
Mühendisliği Bölümü

Eşref Adalı
İTÜ Bilgisayar
Mühendisliği Bölümü

{senturkfi, [adali](mailto:adali@itu.edu.tr)}@itu.edu.tr

Özetçe

Bu çalışma ile Türkçe metinler için kural tabanlı bir bul ve değiştir algoritması geliştirilmiş ve uygulanmıştır. Çalışmanın gerekliliği Türkçenin bitişken bir dil olmasından kaynaklanmaktadır. Mevcut bul ve değiştir yazılımları bükümlü diller için kullanılabilmektedir. Bilindiği gibi bitişken diller, dolayısıyla Türkçe çok sayıda ek almaktadır. Bu nedenle Türkçe için ekleri de göz önüne alan ve buna bağlı olarak bul ve değiştir işlevini yerine getiren bir yazılıma gerek duyulmuştur.

Türkçenin eklemeli dil olması nedeniyle bulunacak sözcüğün biçimbilimsel çözümlemesinin yapılmasının, ayrıca yerleştirilecek sözcüğün bu çözümlemeye uygun olarak yapılandırılmasının sağlanması gerekmektedir. Tüm bu nedenlerden dolayı böyle bir çalışmanın yapılmasına karar verilmiştir. Hem bul hem de değiştir kısımları için gereken algoritmalar da tasarlanmış ve gerçekleştirilmiştir.

Abstract

Competent Find And Replace

In this work a rule based find and replace functions for texts in Turkish are investigated. The necessity for this work is coming from the fact that Turkish is an agglutinative language. The available find and replace functions can be used for fusional languages. It is very well known that the languages which are agglutinative are taking a lot of annexes. Turkish, which is one of the agglutinative languages, is known that it is one of the languages which is taking a lot of annexes. Therefore there is a need for the software which is taking the annexes in the consideration when is doing find and replace operations.

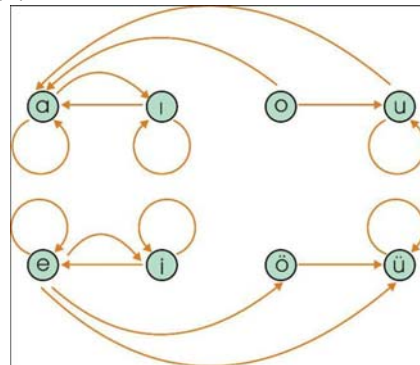
Since Turkish is an agglutinative language, firstly the morphological analysis for word which is tried to be found must be done and the word which will replace it must be configured in appropriate form. The reasons which are mentioned above gave the motivation for this work. The algorithms are designed and implemented both for find and replace function parts in this work.

1. Giriş

Bul ve değiştir işlevleri çoğunlukla birlikte kullanılan işlevlerdir. Bul işlevi sadece aranan metin içindeki bir sözcük ya da harf dizisini bulma amaçlı tek başına kullanılabilirken, değiştir işlevi bulunan sözcüğün yerine başka sözcüğü yerleştirirken kullanılır. Günümüzde yaygın olarak kullanılan yazım programları içinde bulunan "Bul ve Değiştir" işlevleri genellikle bükümlü diller için geliştirilmiş olduğundan Türkçe gibi bitişken diller için yeterli değildir. Ek zengini olan Türkçede bul ve değiştir dediğimiz birimler, aslında sözcüğün kökü ve ekleridir. Ekler, Türkçenin dilbilgisi kurallarına göre şekillendiğinden, değiştirme işlemi sırasında eklerin yeni köke uygun olarak biçimlenmesi gerekir. Bu kuralların en önde gelenleri büyük ve küçük ünlü uyumlarıdır. Bu kuralların Sonlu Durum Makineleri (SDM) Şekil-1'de gösterilmiştir. Aşağıda bul ve değiştir işlevinin kullanılması sırasında Türkçeye özgü büyük ve küçük ünlü uyumlarıyla ilgili çıkan sorunların bir kısmı gösterilmiştir.

Örneğin;

"Kedilerimizden siyah olanlar kaçtı." cümlesinde "kedi" sözcüğünün yerine "tavşan" sözcüğü konulmak istenildiğinde mevcut yazılımlarda değişim "tavşanlarımızdan siyah olanlar kaçtı." şeklinde gerçekleşmektedir. Bu çalışma ile, değişimin "Tavşanlarımızdan siyah olanlar kaçtı." şeklinde yapılması sağlanmıştır.



Şekil-1: Büyük ve küçük ünlü uyumu [1]

Bul ve deęiřtir iřlevi sırasında karřılařılan bir bařka sorun s6zcük sonundaki 6ns6zlerin deęiřmesidir.

6rnek;in;

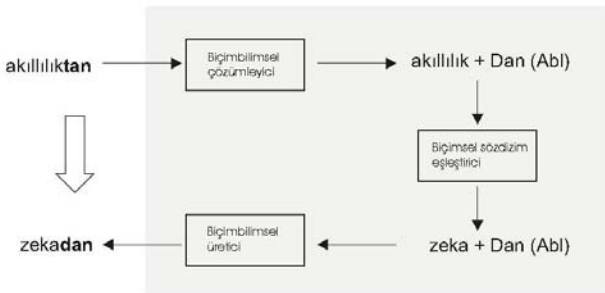
Aranan s6zcük “kitap” olduęunda “Kitabımı getirseydin 6alıřırdık.” c6mlesinde bulma ger6ekleřemez. Bu nedenle “kitap” s6zc6ę6n6n son 6ns6z6n6n yumuřamıř hali de aramaya katılmıřtır. Karřılařılan bařka bir sorun da 6zellikle Arap6adan dilimize girmiř olan s6zc6klerde karřılařılan ses d6řmesi olayıdır.

6rnek;in;

Aranan s6zcük “oęul” ise “oęlumun g6mleęi kırmızı” s6zcük 6beęinde bulma ger6ekleřmez. Bu nedenle “oęul” s6zc6ę6 gibi ses d6řmesine uęrayan s6zc6klerde s6zc6ę6n ses d6řmesine uęramıř hali de aramaya katılmıřtır.

2. Benzer 6alıřmalar

T6rk6e dıřındaki bazı dillerde bul ve deęiřtir iřlevi doęru sonu6 verebilmektedir. 6rnek;in; İngilizce i6in sayısı beři ařmayan bazı 6zel durumlar dıřında, bulma ve doęru deęiřtirme yapılabilir. Buna karřın T6rk6e dili i6in bazı 66z6mler sunulmaya 6alıřıldıysa da kullanılabilir bir 6rnek bulunmamaktadır. O. Bilgin, 6. 6etinoęlu ve K. Oflazer'in tarafından hazırlanan ancak yayınlanmayan “Efficient Find and Replace in Agglutinative Languages: The Case of Turkish” isimli 6alıřmalarında, bitiřken dillerde kullanılmak 6zere T6rk6e i6in 6zelleřtirilmiř etkili bir bul ve deęiřtir iřlevi 6nerilmiřtir. Bu 6alıřma, bulunacak olan s6zc6ę6 bi6imbirimsel 66z6mleyicide iřleyerek elde edilen sonucu kullanarak yerine konması istenen s6zc6ę6 yapılandırmaya dayanmaktadır. Bu 6alıřmada dolabını s6zc6ę6ndeki “dolap” yerine “masa” s6zc6ę6 konulmak istendięinde ilk olarak “dolabını” s6zc6ę6 bi6imbirimsel 66z6mleyicide 66z6mlendięinde “dolap+Noun+A3sg+P2sg+Acc (senin dolabını)” ve “dolap+Noun+A3sg+P3sg+Acc (onun dolabını)” gibi iki 66z6m elde edilmektedir. Bu 66z6mlerden elde edilen ek ve k6k bilgileri “masa” s6zc6ę6 i6in iřlenerek “masa+Noun+A3sg+P2sg+Acc” → “masanı” ve “masa+Noun+A3sg+P3sg+Acc” → “masasını” 66z6mlerini 6retilir. B6ylece, bařarılı bir řekilde bul ve deęiřtir iřlevi ger6ekleřtirilmiř olur. řekil-2’de “akıllılık” s6zc6ę6 yerine “zeka” s6zc6ę6n6n konulması ve ge6ilen ařamalar g6sterilmiřtir [3].



řekil-2: “akıllılık” s6zc6ę6n6n “zeka” s6zc6ę6 ile deęiřtirilmesi

3. Kural Belirleme

Bu 6alıřma kapsamında bul ve deęiřtir iřlevlerine 66z6m 6retmeye 6alıřılmıřtır. Bu nedenle s6zc6klerin k6k ve eklerinin bi6imbirimsel dizilimlerinin bir kurala oturtulması saęlanmıřtır. Bu baęlamda t6m 6ekim eklerinin ve bazı yapım eklerinin de aralarında bulunduęu bir ek k6mesi i6in sonlu durum makineleri 6ıkarılmıřtır ve bu sonlu durum makineleri kullanarak deęiřtir iřlevinin doęru 6alıřabilmesi i6in zemin oluřturulmuřtur. Bu b6l6mde ad soylu s6zc6klerin 6ekim ekleri ve ek eylem ekleri,

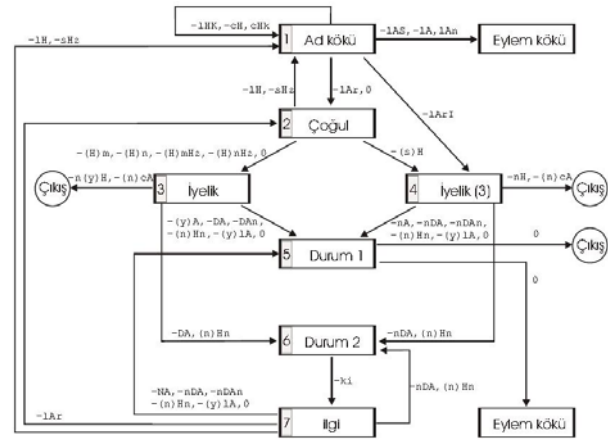
eylem soylu s6zc6klerin 6ekim ve zaman ekleri ve bunlara bazı yapım ekleri de eklenerek t6m s6zc6kler i6in sonlu durum makineleri oluřturulmuřtur. T6m bu sonlu durum makinelerinin birlikte 6alıřabilmeleri i6in 66z6m bulunmuřtur. 6alıřma kapsamında incelenen ek sayısı toplamı 115 tir. Bu eklerin hangi t6r s6zc6klerle nasıl hareket ettikleri bu b6l6mde g6sterilmiřtir.

3.1. Ad 6ekim Ekleri

Ad 6ekim ekleri ad soylu s6zc6klere eklenerek onlara t6mcede g6rev ve anlam kazandıran eklerdir. 6oęul, iyelik, ilgi ve durum ekleri ad 6ekim eki olarak bilinir.

6rnek;in;

“6ocuklarını” s6zc6ę6nde “6ocuk” s6zc6ę6 sırasıyla “lar” 6oęul, “ın” iyelik ve “ını” ilgi eki almıřtır. řekil-3’te Ad 6ekim eklerinin ad k6k6ne nasıl eklendiklerinin kuralları g6r6lmektedir. Bu řekle baęlı kalarak ad 6ekim ekleri i6in sonlu durum makinesi yapılmıřtır. 6alıřmanın kapsamında ad k6k6ne eklenen ekler soldan saęa doęru olmak 6zere incelenmiřtir.

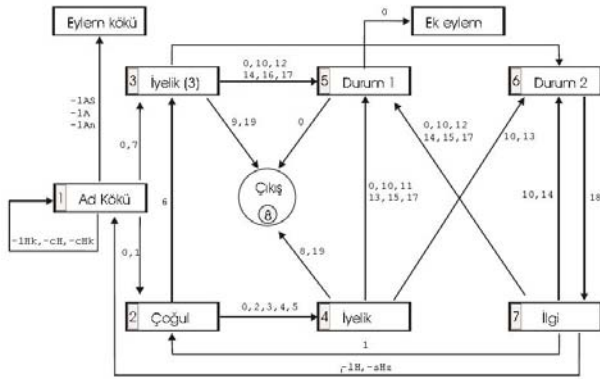


řekil-3: Ad 6ekim ekleri soldan saęa sonlu durum makinesi [2][4]

Bu sonlu durum makinesinde herhangi ad soylu bir s6zc6ę6 k6k ve eklerine ayırabilmemiz m6mk6nd6r. Ad soylu s6zc6klere eklenen ve ad 6ekim eki olarak tanımlanan 19 ek ve 6 yapım eki i6in řekil-3’teki sonlu durum makinesi 66z6m saęlamaktadır. Bu řekli daha matematiksel bir hale d6n6řt6rerek bilgisayar sistemlerinin anlayabileceęi bir řekle getirilmesi gerekmektedir. Bunun i6in 6ncelikle bu 19 ad 6ekim ekinin numaralandırması ve ona g6re řeklin yeniden yorumlanması saęlanmıřtır. “0” numaralı ge6iřler boř ge6iřleri ifade etmektedir. řekil-4’te, 6izelge-1’deki ek katarlarının yerine ek numaralarının geliři g6sterilmiřtir (Yapım ekleri katar halinde g6sterilmiřtir).

6izelge-1: Ad 6ekim Eklerinin Numaralandırılması [2]

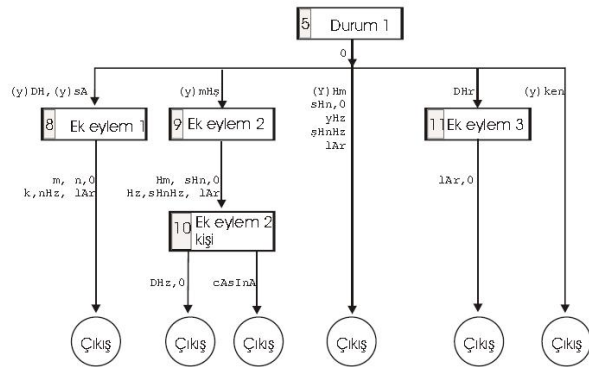
Ek No:	Ek	Açıklama	Örnek
1	-lAr	Çoğul	kedi-ler
2	-(H)m	1. tekil kişi iyelik	kedi-m
3	-(H)mHz	1. çoğul kişi iyelik	kedi-miz
4	-(H)n	2. tekil kişi iyelik	kedi-n
5	-(H)mHz	2. çoğul kişi iyelik	kedi-niz
6	-(s)H	3. tekil kişi iyelik	kedi-si
7	-lArı	1. çoğul kişi iyelik	kedi-leri
8	-(y)H	-i hali	kedi-yi
9	-nH	-i hali (3.t.k. iyelikten sonra)	kedi-ni
10	-(n)Hn	Tamlama	kedi-nin
11	-(y)A	-e hali	kedi-ye
12	-nA	-e hali (3.t.k. iyelikten sonra)	kedi-ne
13	-DA	-de hali	kedi-de
14	-nDA	-de hali (3.t.k. iyelikten sonra)	kedi-nde
15	-Dan	-den hali	kedi-den
16	-nDAn	-den hali (3.t.k. iyelikten sonra)	kedi-nden
17	-(y)lA	birliklilik	kedi-yle
18	-ki	İlgi	kedi-de-ki
19	-(n)cA	Görelilik	kedi-ce



Şekil-4: Numaralandırılmış ad çekim ekleri soldan sağa sonlu durum makinesi

3.2. Ek Eylem Ekleri

Tıpkı eylemlerde olduğu gibi adlara da ek eylem ekleri ile zaman ve kişi anlamları yüklenebilmektedir. Bu nedenle bu eklerin de incelenmesine gerek duyulmuştur. Ek eylem eklerinin ad soylu sözcüklere eklenme kuralları Şekil-5'te gösterilmiştir.

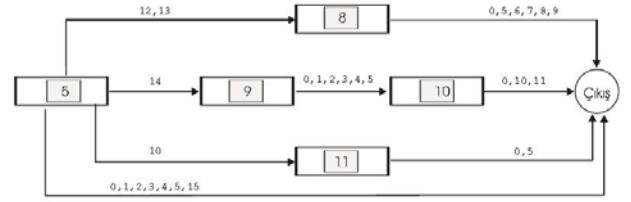


Şekil-5: Ek eylem ekleri soldan sağa sonlu durum makinesi [2][4]

Ad çekim eklerinde yapıldığı gibi ek eylem ekleri için numaralandırılmış sonlu durum makinesi Çizelge 2'deki ek numaraları yardımıyla Şekil-6'da oluşturulmuştur.

Çizelge-2: Ek Eylem Eklerinin Numaralandırılması [2]

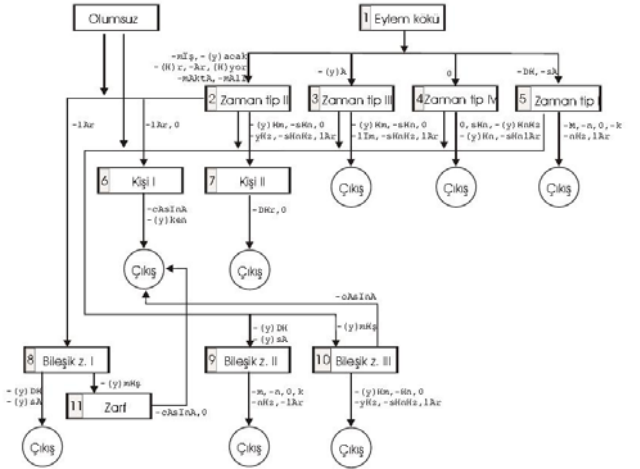
Ek No:	Ek	Açıklama	Örnek
1	-(y)Hm	1. tekil kişi	evde-yim
2	-sHn	2. tekil kişi	evde-sin
3	-(y)Hz	1. çoğul kişi	evde-yiz
4	-sHnHz	2. çoğul kişi	evde-siniz
5	-lAr	3. çoğul kişi	evde-ler
6	-m	1. tekil kişi ((y)DH ve (y)sA eklerinden sonra)	evdeyse-m
7	-n	2. tekil kişi ((y)DH ve (y)sA eklerinden sonra)	evdeyse-n
8	-k	1. çoğul kişi ((y)DH ve (y)sA eklerinden sonra)	evdeyse-k
9	-nHz	2. çoğul kişi ((y)DH ve (y)sA eklerinden sonra)	evdeyse-niz
10	-DHr	çevrik kip	evde-dür
11	-cAsInA	tarz zarfı	ev-cesine
12	-(y)DH	di'li geçmiş zaman	evde-ydi
13	-(y)sA	Dilek-şart kipi	evde-yse
14	-(y)mHş	miş'li geçmiş zaman	evde-ymiş
15	-(y)ken	zaman zarfı	evde-yken



Şekil-6: Numaralandırılmış ek eylem ekleri soldan sağa SDM

3.3. Eylem Zaman Ekleri

Eylem soylu sözcüklere zaman ve kişi anlamları katan eklere eylem zaman ekleri denir. Şekil-7'de eylem zaman eklerinin eylem soylu bir sözcüğe hangi kurallara göre eklenebileceği görülmektedir.

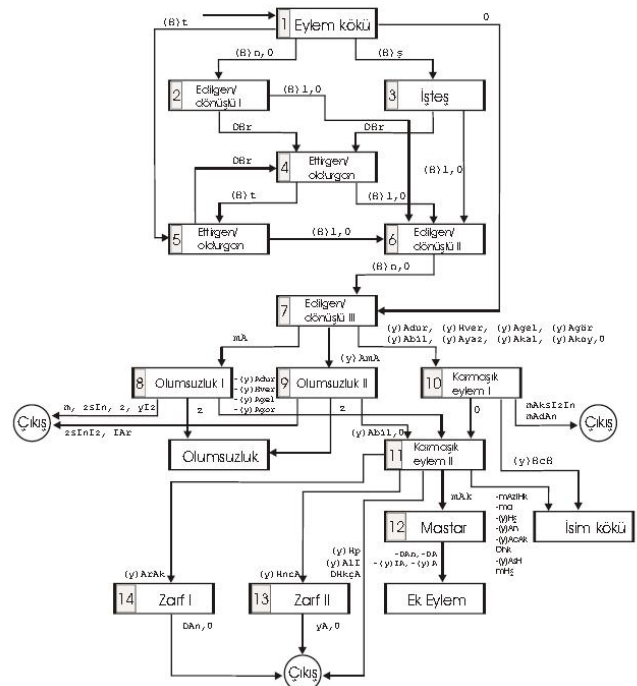


Şekil-7: Eylem zaman ekleri soldan sağa SDM [2][4]

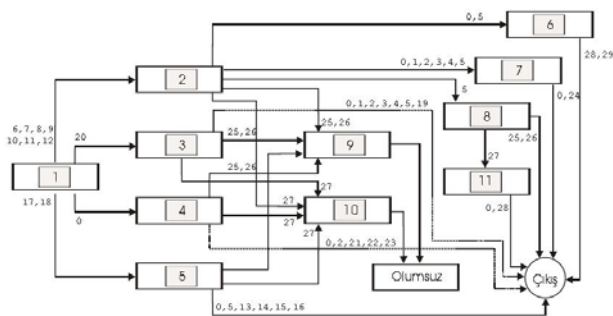
Çizelge 3'teki eklerin numaraları yardımıyla eylem zaman ekleri için de tıpkı önceki 2 ek grubunda yapıldığı gibi oluşturulmuş numaralandırılmış solda sağa sonlu durum makinesi Şekil-8'de görülmektedir.

Çizelge-3: Eylem Zaman Eklerinin Numaralandırılması [2]

Ek No:	Ek	Açıklama	Örnek
1	-(y)Hm	1. tekil kişi	gel-iyor-um
2	-sHn	2. tekil kişi	gel-iyor-sun
3	-(y)Hz	1. çoğul kişi	gel-iyor-uz
4	-sHnHz	2. çoğul kişi	gel-iyor-sumuz
5	-lAr	3. çoğul kişi	gel-iyor-lar
6	-mHş	miş'li geçmiş zaman	gel-miş
7	-(y)AcAk	gelecek zaman	gel-miş
8	-(H)r	geniş zaman	gel-ir
9	-Ar	geniş zaman	gid-er
10	-(H)yor	şimdiki zaman	gel-iyor
11	-mAktA	sürerlilik	gel-mekte
12	-mAll	gereklilik	gel-meli
13	-m	1. tekil kişi	gel-di-m
14	-n	2. tekil kişi	gel-di-n
15	-k	1. çoğul kişi	gel-di-k
16	-nHz	2. çoğul kişi	gel-di-niz
17	-DH	di'li geçmiş zaman	gel-di
18	-sA	dilek-şart kipi	gel-se
19	-lIm	1. çoğul kişi	gel-e-lim
20	-(y)A	istek kipi	gel-e
21	-(y)HnHz	2. çoğul kişi	gel-iniz
22	-(y)Hm	2. tekil kişi	gel-in
23	-sHnAr	3. çoğul kişi	gel-sinler
24	-DHr	çevrik kip	gel-miş-ler-dir
25	-(y)DH	hikaye bileşik zaman	gel-miş-ti
26	-(y)sA	şart bileşik zaman	gel-miş-se
27	-(y)mHş	rivayet bileşik zaman	gel-meli-ymiş
28	-cAsInA	tarz zarfı	gel-miş-cesine
29	-(y)ken	zaman zarfı	gel-miş-ken



Şekil-9: Eylem çekim ekleri soldan sağa SDM [2][4]



Şekil-8: Numaralandırılmış eylem zaman ekleri soldan sağa SDM

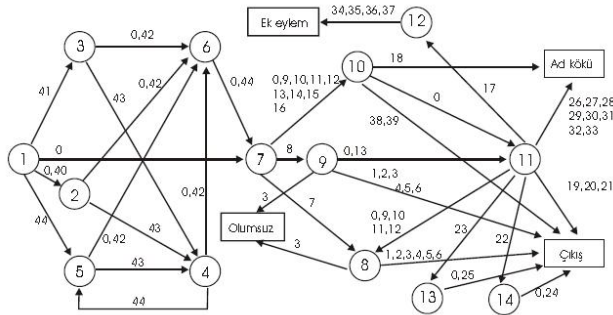
3.4. Eylem Çekim Ekləri

Eylem Zaman ekleri bölümünde eylem soylu sözcüklere eklenen kişi ve zaman ekleri açıklanmıştır. Bu eklerin dışında da eylem soylu sözcüklere eklenen çekim ekleri vardır. Bu bölümde Eylem çekim ekleri olarak adlandırdığımız ekler tanıtılacaktır. Şekil-9'da eylem çekim eklerinin eylem soylu bir sözcüğe hangi kurallara göre eklenebileceği görülmektedir.

Çizelge-4: Eylem Çekim Eklerinin Numaralandırılması [2]

Ek No:	Ek	Açıklama	Örnek
1	-m	1. tekil kişi	gel-me-m
2	-sln	2. tekil kişi	gel-me-sln
3	-z	3. tekil kişi	gel-me-z
4	-ylyz	1. çoğul kişi	gel-me-ylyz
5	-zslnyz	2. çoğul kişi	gel-me-zslnyz
6	-zlar	3. çoğul kişi	gel-me-zler
7	-mA	Olumsuzluk	gel-me
8	-(y)AmA	Olumsuzluk	gel-eme
9	-(y)Adur	sürerlik bileşik eylem kipi	gel-edur
10	-(y)Hver	tezlik bileşik eylem kipi	gel-iver
11	-(y)Agel	sürerlik bileşik eylem kipi	ol-ager
12	-(y)Agör	sürerlik bileşik eylem kipi	gel-egör
13	-(y)Abil	yeterlik bileşik eylem kipi	gel-ebil
14	-(y)Ayaz	yaklaşma bileşik eylem kipi	gel-eyaz
15	-(y)Akal	sürerlik bileşik eylem kipi	don-akal
16	-(y)Akoy	sürerlik bileşik eylem kipi	al-akoy
17	-mAk	Master	gel-mek
18	-(y)HeH	görev eki	gör-üdü
19	-(y)Hp	zarf eki	gel-üp
20	-(y)Alı	zarf eki	gel-elı
21	-DHkeA	zarf eki	gel-dike
22	-(y)ArAk	zarf eki	gel-erek
23	-(y)HneA	zarf eki	gel-ince
24	-DAñ	zarf eki	gör-erek-ten
25	-yA	zarf eki	gel-ince-ye
26	-(y)An	şafat fiil	gel-en
27	-(y)AeAk	fiilden isim yapma eki	gel-ecek
28	-(y)Ası	şafat fiil	gel-esi
29	-DHk	şafat fiil	gönd-duk
30	-mHş	şafat fiil	gel-mış
31	-mAzlık	fiilden isim yapma eki	gel-mezlik
32	-mA	master eki	gel-me
33	-(y)Hş	fiilden isim yapma eki	gel-iş
34	-DAn	-den hali	gel-me-den
35	-DA	-de hali	gel-me-de
36	-(y)LA	Birliklik	gel-me-yle
37	-(y)A	-e hali	gel-me-ye
38	-mAksızın	zarf eki	gel-meksizin
39	-mAdAn	zarf eki	gel-medan
40	-(H)n	Edilgen dönüşü	gör-ün
41	-(H)ş	İştes	gör-üş
42	-(H)l	Edilgen dönüşü	gör-ül
43	-DHr	Oldurgan	ol-dur
44	-(H)t	Ettirgen	oku-t

Çizelge 4'teki eklerin numaraları yardımıyla eylem çekim ekleri için de tıpkı önceki 3 ek grubunda yapıldığı gibi oluşturulmuş numaralandırılmış solda sağa sonlu durum makinesi Şekil-10'da görülmektedir.



Şekil-10: Numaralandırılmış eylem çekim ekleri soldan sağa SDM

3.5. Tüm Sonlu Durum Makinelerinin Ortak Sonlu Durum Makinesinde Toplanması

Şekil-4, Şekil-6, Şekil-8 ve Şekil-10 soldan sağa doğru numaralandırılmış sonlu durum makinelerinin tümünde birbirlerine geçişler mevcuttur.

Örneğin:

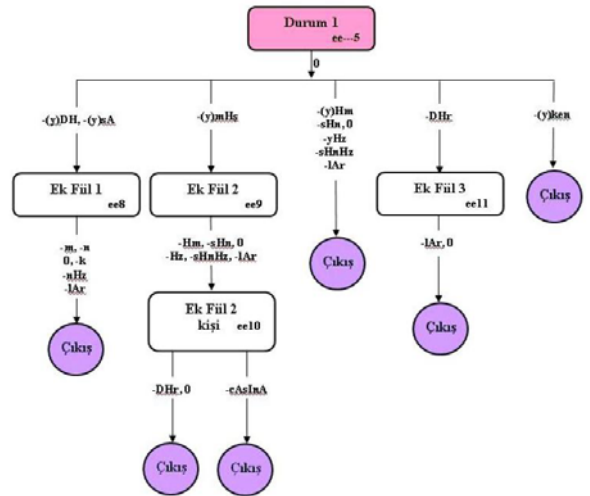
Şekil-3 Ad kökü sonlu durum makinesinin 5 numaralı durumunda Ek Eylem makinesine geçiş vardır. Benzer şekilde Şekil-9 Eylem Çekim ekleri makinesinde 8 ve 9 numaralı durumlardan Eylem Zaman ekleri sonlu durum makinesinin OLUMSUZ durumuna, 10 ve 11 numaralı durumlardan Ad Kökü sonlu durum makinesine ve 12 numaralı durumdan da Ek Eylem sonlu durum makinesine geçişler vardır. Tüm bu geçişlerin tek bir sonlu durum makinesinde bir araya alınmasının ardından, 0 (sıfır) boş geçişlerin elenmesi gerekmektedir. Tüm eklerin bir arada kullanılabilmesi amacıyla her bir ek grubunda 1'den başlanarak numaralandırılan ekler için teklifi sağlamak amacıyla 1 ile 115 arasında numaralar verilmiştir. Bu numaraların hangi eklerle denk geldiği EK A.'da görülmektedir. [1-44] ek numaraları Eylem Çekim Ekleri, [45-59] ek numaraları Ek Eylem Ekleri, [60-78] ek numaraları Ad Kökü Çekim Ekleri, [79-107] ek numaraları Eylem Zaman Ekleri ve [108-115] ek numaraları Yapım Ekleri olarak sınıflandırılmıştır.

Ortak olarak oluşturulmuş sonlu durum makinesinin burada gösterimi boyutundan dolayı uygun değildir. Bu nedenle gösterilmemiştir.

Ayrıca boş geçişlerin de nasıl elendiğini küçük bir örnek kümesi ile göstermekte fayda vardır. Bu sayede son oluşturulan büyük sonlu durum makinesinde boş geçişlerin nasıl elendiği ile ilgili fikir sahibi olunabilir.

Boş Geçişlerin Elenmesi

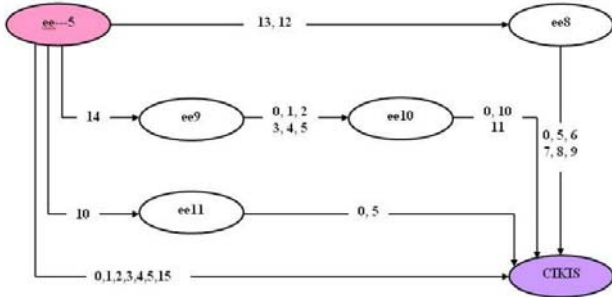
Şekil-11 ve Çizelge 5 yardımıyla Şekil-12 oluşturulmuştur. Şekil-12 de 0 (sıfır) boş geçişlerin elenmesi için birkaç işlem yapılacaktır. Tüm işlemler aşağıda adım adım gösterilmiştir. İlk olarak hangi durumlara hiç bir işlem yapmadan ulaşılabilir ona bakılır ve A kümesine bu durumlar eklenir. Şekil-12'de A kümesinde ee5 durumu bulunmaktadır. A kümesinde hangi ekler ile hangi kümelere gidileceğini gösteren aşağıdaki işlemler yapılmıştır.



Şekil-11: Ek eylem ekleri soldan sağa sonlu durum makinesi[2][4]

Çizelge-5: Ek Eylem Eklerinin Numaralandırılması [2]

EK NUMARASI	EK	AÇIKLAMA	ÖRNEK
1	-(y)Hm	1. tekil kişi	evde-yim
2	-sHn	2. tekil kişi	evde-sin
3	-(y)Hz	1. çoğul kişi	evde-yiz
4	sHnHz	2. çoğul kişi	evde-siniz
5	-lAr	3. çoğul kişi	evde-ler
6	-m	1. tekil kişi ((y)DH ve (y)sA elderinden sonra)	evdeyse-m
7	-n	2. tekil kişi ((y)DH ve (y)sA elderinden sonra)	evdeyse-n
8	-k	1. çoğul kişi ((y)DH ve (y)sA elderinden sonra)	evdeyse-k
9	-nHz	2. çoğul kişi ((y)DH ve (y)sA elderinden sonra)	evdeyse-niz
10	-DHr	çevrik kâp	evde-dir
11	-cAşnA	tarz zarfı	ev-cesine
12	-(y)DH	dâ'li geçmiş zaman	evde-ydi
13	-(y)sA	Dilek-şart kâpi	evde-yse
14	-(y)mHş	miş'li geçmiş zaman	evde-ymiş
15	-(y)ken	zaman zarfı	evde-yken



Şekil-12: Numaralandırılmış ek eylem ekleri soldan sağa SDM

A={ee5}

ee5 durumunda çıkan ekler 1,2,3,4,5,10,12,13,14,15

1 eki ile Çıkış
2 eki ile Çıkış
3 eki ile Çıkış
4 eki ile Çıkış
5 eki ile Çıkış
10 eki ile {ee11} → {ee11,Çıkış}=B
12 eki ile {ee8} → {ee8,Çıkış}=C
13 eki ile C
14 eki ile {ee9} → {ee9,ee10,Çıkış}=D
15 eki ile Çıkış

B={ee11,Çıkış}

ee11 durumunda çıkan ekler 5
5 eki ile Çıkış

C={ee5,Çıkış}

ee8 durumunda çıkan ekler 5,6,7,8,9
5 eki ile Çıkış
6 eki ile Çıkış
7 eki ile Çıkış
8 eki ile Çıkış
9 eki ile Çıkış

D={ee9,ee10,Çıkış}

ee9 durumunda çıkan ekler 1,2,3,4,5
1 eki ile {ee10} → {ee10,Çıkış}=E
2 eki ile E
3 eki ile E
4 eki ile E
5 eki ile E

ee10 durumunda çıkan ekler 10,11

10 eki ile Çıkış
11 eki ile Çıkış

E={ee10,Çıkış}

ee10 durumunda çıkan ekler 10,11
10 eki ile Çıkış
11 eki ile Çıkış

Tüm bu işlemlerin sonucunda aşağıdaki 6 küme oluşmaktadır.

A={ee5}

B={ee11,Çıkış}

C={ee8,Çıkış}

D={ee9,ee10,Çıkış}

E={ee10,Çıkış}

Çıkış

Yukarıda örneklendirdiğimiz boş geçişlerin elenmesi tüm sonlu durum makinelerinin bir arada olduğu sonlu durum makinesi için yapıldığında 36 kümeden oluşan bir yapıya ulaşılmaktadır. Durumların kolay anlaşılması için bazı kısaltmalara başvurulmuştur. Bunlar ; ez: Eylem Zaman, ik: Ad (isim) kökü ve ee: Ek eylem dir.

A={1,2,6,7,10,11,eZ4}

B={2,6,7,10,11,eZ4}

C={3,6,7,10,11,eZ4}

D={5,6,7,10,11,eZ4}

E={4,6,7,10,11,eZ4}

F={6,7,10,11,eZ4}

G={7,10,11,eZ4}

H={8,11,eZ4}

I={9,11,eZ4}

J={10,11,eZ4}

K={ik1,ik2,ik3,ik4,ik5,ee}

L={ez2,eZ7,eZ6,OLUMSUZ,ARA}

M={ez3, ARA}

MA={ez5, ARA}

N={14}

O={13}

P={12}

R={OLUMSUZ,ARA}

S={11,eZ4}

T={ik1,ik3,ik5,ee}

U={ik4,ik5,ee}

V={ik3,ik5,ee}

Y={ik5,ee}

Z={ik5,ik6,ee}

A1={ee8}

B1={ee9,ee10}

C1={ee11}

D1={ez6,eZ7,eZ8}

E1={ez7}

F1={ez8}

G1={ez9}

H1={ez10}

I1={ee}

J1={ik7,ik5,ee}

K1={ee10}

ÇIKIŞ={}

A kümesi başlangıç kümesi iken ÇIKIŞ kümesi sonlanmayı ifade etmektedir. Eklerin çözülmesi aşamasında tüm durumlarda sonlanma olabilmektedir. Ek katarının parçalanması hangi durumda biterse orda ek çözülmesi sonlanmış demektir. Bu bilgiler ışığında yazılımda kullanılan geçiş kümeleri oluşturulmuştur.

4. Yetkin Bul ve Değiştir

Türkçe için karşılaşılan birçok bulma ve değiştirme sorunu vardır. Bu sorunların birçoğuna değinilerek çözüm sunulmaya çalışılmıştır.

4.1 Bulmada Karşılaşılan Sorunlar

Türkçe için karşılaşılan birçok bulma ve değiştirme sorunu vardır. Bu sorunların bazılarını ve çözümlerine aşağıda değinilmiştir.

- Türkçeye yabancı dillerden girmiş bazı sözcüklerde ünlü düşmesi yaşanır.

Örneğin:

“Ayşe oğlunu gezmeye götürdü.” tümcesinde “oğul” sözcüğü arandığında hali hazırdaki bul işlevleri “oğlunu” sözcüğünü yakalayamamaktadır. Bu nedenle ünlü düşmesi yaşayan sözcükler bir dosyada tutularak, bul denen sözcük bu dosyada var ise ünlü düşmesi yaşamış halinin de aramaya katılması sağlanmıştır.

- Son harfi “p, ç, t, k” sert süreksiz ünsüzü olan sözcükler ünsüz yumuşamasına uğradıklarında değişirler. Aramaya hem değişmemiş hem de değişmiş hallerinin katılması sağlanmıştır.

Örneğin:

“Kemal kitabını kitaplıktan aldı.” tümcesinde “kitap” sözcüğü arandığında sözcüğün yumuşamaya uğradığı “kitab” sözcüğü de aramaya katılarak, hem “kitap” hem de “kitab” aranmalıdır. Böylece hem “kitabını” hem de “kitaplıktan” sözcükleri yakalanarak bulunmalıdır.

- Son harfi “p, ç, t, k” sert süreksiz ünsüzü olduğu halde ünsüz yumuşamasına uğramayan sözcükler de vardır. Bu özelliğe sahip sözcükler bir dosyada tutularak, yumuşamaya uğratmadan olduğu gibi aranması sağlanmıştır.

Örneğin:

“aşk” sözcüğü ünsüz yumuşaması kurallarına göre ünlü bir harf ile başlayan ek aldığında sondaki “k” ünsüzünün “g” ünsüzüne dönüşmesi beklenir, fakat değişme olmaz. Bu nedenle bu özelliğe sahip sözcükler bir dosyada tutularak, bu örnek için hem “aşk” hem de “aşg” katarlarının aramaya katılıp verimin azalması engellenmiştir.

- Bazı eylem köklü sözcüklere “yor” eki eklendiğinde eylem soylu sözcükte değişimler olur. Bu değişimler göz önüne alınarak bir arama yapılması sağlanmıştır.

Örneğin:

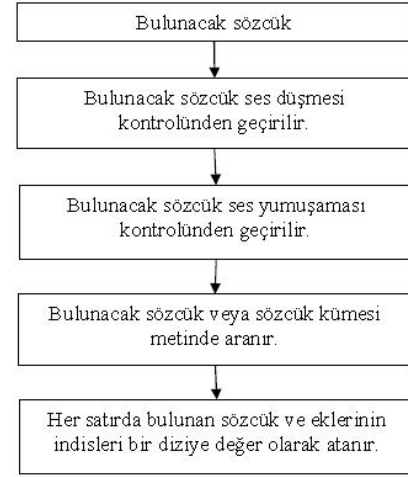
“anla” eylem soylu sözcüğüne “yor” eki eklendiğinde “anla+yor” yerine “anlıyor” haline dönüş olur. “Kenan her söyleneni anlıyor.” tümcesinde “anla” sözcüğü arandığında bulunamıyor. Bu nedenle bu şartları sağlayan eylem soylu sözcükler için özel bir çözümleme yapılmıştır.

- “de” ve “ye” eylem soylu sözcüklerinde “ecek” eki eklendiğinde eylem soylu sözcükte değişimler olur. Bu değişimler göz önüne alınarak bir arama yapılması sağlanmıştır.

Örneğin:

“de + ecek” → “diyecek”

“ye + ecek” → “yiyecek”



Şekil-13: Bul işlevi akış şeması

Şekil-13'te görüldüğü üzere bul işlevinde aranan sözcük bulunmadan önce sorun yaratan durumların kontrolünün yapıldığı kontrollerden geçildikten sonra bulunan sözcüğün ardındaki ek katarı bir çözümleyiciden geçirilerek her bir ek katarı için ek numaralarından oluşan bir dizi oluşturulur.

4.2 Değiştirmede Karşılaşılan Sorunlar

Değiştir işlevinde de tıpkı bul işlevinde olduğu gibi dile özgü sorunlarla karşılaşmaktadır. Bu sorunların başlıca nedenleri Türkçenin sondan eklemeli bir dil olmasından ve eklerinin neyin ardına eklendiğine göre değişmesinden kaynaklanmaktadır.

Örneğin:

“Özkan aldığı kitaplarından bazılarını yenileriyle değiştirdi.” tümcesinde “kitap” sözcüğü bulunup yerine “defter” sözcüğü konulmaya çalışıldığında, tümce “Özkan aldığı defterlerinden bazılarını yenileriyle değiştirdi.” halini almaktadır. “defterlerinden” sözcüğü Türkçede anlamlı bir sözcük değildir. Bunun yerine “defterlerinden” sözcüğü olmalıdır. Bu ve benzeri sorunları çözebilmek için yazılımda çözümler üretilmiştir. Değiştir işlevinde karşılaşılan sorunlar ve çözümleri şöyledir. Türkçede sözcükler ve ekleri arasında nasıl bir ilişki olduğu tanımlanmış ve buna göre eklerin yapılandırılması sağlanmıştır.

Örneğin:

“kedilerimizden” sözcüğünde “kedi” sözcüğü yerine “tavşan” değişikliği yapılmaya çalışıldığında “kedi” sözcüğünün ardında hangi ekler olduğuna bakılarak, “ler” çoğul, “imiz” 1. çoğul iyelik ve “den” ayrılma durum eki çözümlemesi yapılması sağlanmıştır. Bu çözümlemenin ardından “tavşan” sözcüğüne ilgili ekler “tavşan” sözcüğünün sahip olduğu ses özelliklerine göre eklenmiş ve “tavşanlarımızdan” son sözcüğünün yapılandırılması sağlanmıştır.

- Eğer Bul işlevinde değinilen ünlü düşmesi olan sözcüklerden biri bulunmuş sözcüğün yerine konacak sözcük olarak seçildiyse olası değişiklikler göz önüne alınarak değiştir işlevinin yapılması sağlanmıştır.

Örneğin:

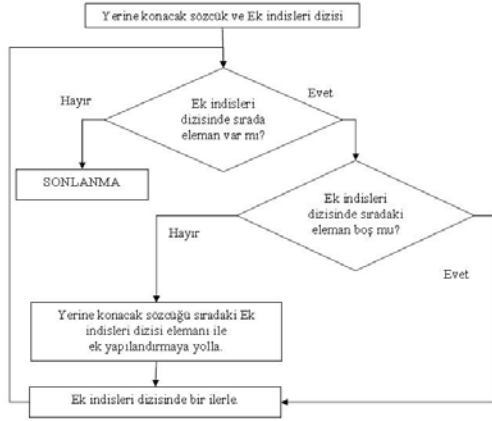
“Ahmet Bey kızının okuluna hiç gitmedi.” tümcesinde “kız” yerine “oğul” değişikliği yapılmaya çalışıldığında, “oğul” sözcüğünün ses düşmesi yaşayan sözcükler dosyasında olduğu ve ne durumlarda ses düşmesi yaşadığı bilgileri ışığında değişiklik yapılarak tümcenin “Ahmet Bey oğlunun okuluna hiç gitmedi.” halini alması sağlanmıştır.

- Son harfi “p, ç, t, k” ünsüzlerinden biri olan sözcük bulunan sözcük yerine konacak sözcük olarak seçildiyse; ardına eklenen eklerin bazı özelliklerine göre bu sözcüklerde oluşabilecek değişiklikler göz önüne alınarak çözümleme yapılması sağlanmıştır.

Örneğin:

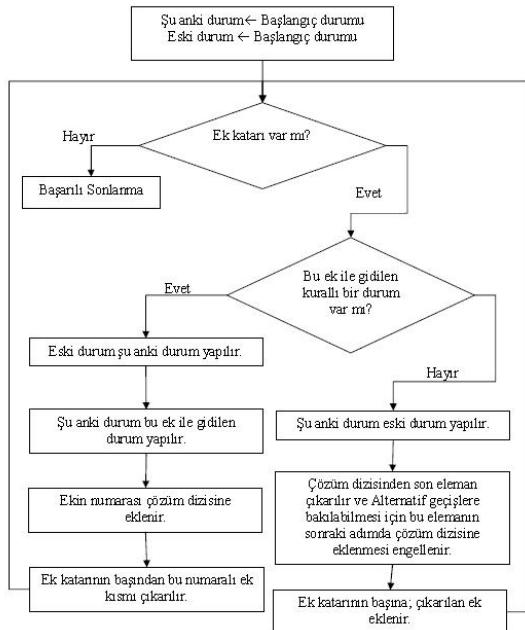
“Özge defterini Selim’in defteriyle karıştırınca olanlar oldu.” Tümcisinde “defter” yerine “kitap” konulmaya çalışıldığında “kitap” sözcüğünün son harfinin “p, ç, t, k” harflerinden biri olması ve ek alırken aldığı ekin ünlü ile başlaması nedeniyle sözcüğün son ünsüzü yumuşamaya uğrar. Bu nedenle tümcenin “Özge kitabını Selim’in kitabıyla karıştırınca olanlar oldu.” halini alması sağlanmıştır.

Şekil-14’te görüldüğü üzere bulunacak sözcüğün ardındaki ek katarları çözümlendikten sonra yerine konacak sözcüğün ardına ek yapılandırılması akış şemasındaki adımları takip ederek yapılmaktadır.



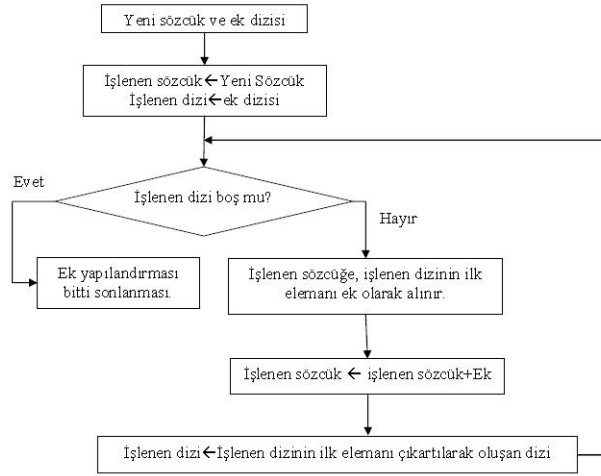
Şekil-14: Değiştir işlevi akış şeması

Hem değiştir hem de bul işlevlerinin nasıl çalıştığı Şekil-13 ve Şekil-14’te görülmektedir. Bulma işlevinin ardından ek katarının çözümlenmesi ve ek dizisinin oluşturulması da önemli bir adımdır. Bu adımın akış şeması da Şekil-15’te görülmektedir.



Şekil-15: Ek dizisi arama algoritması

Şekil-15’teki akış şemasının oluşturduğu ek dizileri kullanılarak yerine konacak sözcüğün ek kısımları yapılandırılır. Bu yapılandırmanın adımları da Şekil-16 da gösterilmiştir.



Şekil-16: Ek yapılandırma işlevi akış şeması

Giriş metni:

osmanlı devleti nedensiz bir şekilde osman beyin ölümü ile osmanlılaşma yolunda osmanlısızlaştırılmaya çalışılmaktadır. osman beyin yerine osmanlı ağası getirilecektir. Ve osmanlılaştıramadıklarımızdan olma durumu ortaya çıkacaktır.

“osman” sözcüğü yerine “sipahi” sözcüğü konulmak istendiğinde aşağıdaki çıkış metni elde edilmektedir..

Çıkış metni:

sipahili devleti nedensiz bir şekilde sipahi beyin ölümü ile sipahilileşme yolunda sipahilisizleştirilmeye çalışılmaktadır. sipahi beyin yerine sipahicik ağası getirilecektir. Ve sipahilileştiremediklerimizden olma durumu ortaya çıkacaktır.

Programı daha büyük dosyalarda da çalıştırarak başarımlı sağlanmıştır. Bunun sonucunda dosyanın büyüklüğünün programın çalışmasında yavaşlamaya neden olmadığı görülmüştür. Metnin içerisine bulunacak olan sözcüğün hem çözümlenebilen hem de çözümlenemeyen şekilleri de konarak programın nasıl bir sonuç vereceği test edilmiştir. Çözümleyicide çözümlenebilen ekleri olan tüm bulunmuş sözcükler çözümlenerek, yerine konacak sözcüğün ardına yapılandırılmıştır ve başarılı bir şekilde kural tabanlı bul ve değiştir sağlanmıştır.

5. Sonuçlar ve Öneriler

Bu çalışmada, Türkçe için özel bir Bul ve Değiştir işlevi tanımlanması yapılmıştır. Bu tanımlama esnasında isim ve eylem soylu sözcükler ve aldıkları eklerin sonlu durum makineleri üzerindeki hareketlerinin kurala bağlanmasıyla kural tabanlı bir değiştir işlevinin başarımlı sağlanabilmiştir. Ayrıca Bul işlevinin de Türkçe için sık sık karşılaştığı bulamama durumu ortadan kaldırılarak başarılı bir bul işlevi yapılandırılmıştır. Metinde aranan sözcük bulunduğu 3 durum söz konusudur. Bunlar;

- Bulunan sözcük ek almamış olabilir. Bu durumda yerine konacak sözcüğe ek yapılandırılması yapılmayacak, olduğu gibi konulacak.
- Bulunan sözcük kurallı bir ek almış olabilir. Bu durumda kurallı ek çözümlenerek, bu çözümlenelerin

verileri ile yeni sözcüğün ekleri yapılandırılacak ve konulması gereken yere konulacak.

- Bulunan sözcük kurallı bir ek almamış olabilir. Bu sorun hangi durumlarda olabilir;
 - Metni yazan kullanıcı sözcüğün eklerini yanlış yazmış olabilir.
 - Ek olarak yapım eki almış bir sözcük aratmış olabilir.

Bu durumda çözümleme yapılamayacağından değişimin olması söz konusu değildir. Bulunan sözcüğün yerine yeni sözcük konulması yapılmadan, eskisi gibi bırakılacaktır.

Bazı yapım ekleri çalışmanın kapsamında incelemeye alınmıştır. Bunların dışında yapım eki içeren sözcüklerde değişim yapılamamaktadır. Bu bir eksiklikten ziyade özellikle engellenmiş bir durumdur. Yapım ekleri sözcüğün anlamında değişikliğe neden olduklarından ek yapılandırılması sırasında doğru bir yapılandırma sonucu vermeyebileceğinden değiştir işlevi için bu çalışma kapsamına alınmamışlardır. Sadece “lHK”, “CH”, “cHK”, “CA”, “lH”, “sHz” isimden isim yapan yapım ekleri ve “lAş”, “lA”, “lAn” addan eylem yapan yapım ekleri çalışmanın kapsamında incelenmiştir.

Kaynakça

[1] **Delibaş, A.**, 2008. Doğal Dil İşleme İle Türkçe Yazım Hatalarının Denetlenmesi, *Yüksek Lisans Tezi*, İ.T.Ü. Fen Bilimleri Enstitüsü, İstanbul.

[2] **Eryiğit, G.**, 2002. Sözlüksüz Köke Ulaşma Yöntemi, *Yüksek Lisans Tezi*, İ.T.Ü. Fen Bilimleri Enstitüsü, İstanbul.

[3] **Bilgin, O., Çetinoğlu Ö., Oflazer K.**, 2005. Efficient Find and Replace in Agglutinative Languages: The Case of Turkish, 1 Mayıs 2008 tarihinde <http://www.hlst.sabanciuniv.edu/archive/patras.pdf> kaynağından alınmıştır.

[4] **Oflazer K.**, 1994. Two-level Description of Turkish Morphology, *Literary and Linguistic Computing*, Vol.9, Number 2.