DEVELOPING A NATURAL LANGUAGE PROCESSING APPLICATION USING GAME PLAYING TECHNIQUES

by Adem Efe Gencer

Submitted to the Department of Computer Engineering in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Engineering

> Boğaziçi University June 2011

Table of Contents

ABSTRACT	iii
INTRODUCTION	1
DETERMINING THE AREA OF RESEARCH	2
PRE-STUDY AND GAME DESIGN	4
Comparison of the Game with Other Games in the Literature	4
Result Submission System and Methods of Submission	4
Time Limit of the Game	5
COLLECTING BASE DATA FOR THE GAME	6
Screenshots from the Initial Base Game	6
Interpretation of the collected Data from Anonymous Participants	10
Using Collected Base Data in Game Award System	10
Badge System	10
Point System	10
BASE DATA RESULTS	11
GAME PLAYING PERIOD	
Collecting User Submission with Known Users	13
Weekly Ranking and Badge Declarations	13
Screenshots from the Game	
CONCLUSION AND FUTURE WORK	19
REFERENCES	20

ABSTRACT

Natural language processing (NLP) is a subcategory of artificial intelligence and linguistics. The ultimate purpose in this area is to let computers understand a natural (human) language and respond it in a similar way. Although the idea is fascinating, there are some hardships in the process. In order to understand a natural language, computers shall be taught the way humans use and learn a language, which is sometimes referred to as an AI-complete problem. Fortunately, on the contrary to computers, some of these tasks could be considered as trivial tasks for humans. Therefore, human-based computation is an extremely useful concept in such situations.

The problem of determining whether a passivized intransitive Turkish verb could be used in the daily language is one of those problems, which is very challenging for computers but trivial for Turkish-speaking humans. Although it is possible to passivize intransitive verbs in a seemingly grammatically correct way, when they are considered within the daily language, it is not possible to use some of them. A human can easily recognize the oddness of the verb within the sentence, but there is no feasible, straightforward algorithm for a computer to do that. For all these reasons, an approach that lets collecting data from humans is followed. In exchange for the data that humans submit, they are given "fun".

In this report, a game, which utilizes from human computation power, with a purpose is introduced. The name of the game is "Dil Cambazı". The game structure and the planned method of interpretation of results are explained. The comparison of the game with other similar games with purposes is investigated. Collected base data is analyzed.

Keywords – Human Computation, Natural Language Processing, Games with a Purpose

1. INTRODUCTION

Natural language processing (NLP) is an appealing topic that is attracting more and more researchers. It lets examination of human language in various ways. Application of human computation via a game is a relatively new approach in this area. The data for determining whether a passivized intransitive Turkish verb could be used in the daily language is collected from users by means of a game. In the end of the game process, the precious data is obtained from the players and in return, the players enjoyed while playing the game.

A game, which utilizes from human computation power, with a purpose is introduced in this report. The game structure and the method of interpretation of results are explained. The comparison of the game with other similar games with purposes is investigated. Collected base data and game data are analyzed. The name of the game was selected as "Dil Cambazı", which is a Turkish idiom meaning that a person who uses the language in an extremely efficient manner.

2. DETERMINING THE AREA OF RESEARCH

Since there are many topics that could be selected to work on in natural language processing, the determination of the area of research was quite a challenging task. There were many alternatives that were considered during the selection process. The selected area had to satisfy certain criteria, such as containing a problem that is easily solvable by people but hard to solve using regular computers and also a unique area that would let the creation of an enjoyable game to collect data from humans.

In order to choose the area, a literature review was performed and the existing ideas were investigated. Especially, the studies of Luis von Ahn were very helpful in this process. The publications that he was a part of were analyzed [1, 2, 3, 4, 5]. Conferences and talks he joined were carefully examined [8, 9]. His idea behind the utilizing human power was studied. The games that were mentioned in von Ahn's papers were tested and pros and cons of those applications were determined. Also the publication of Tunga Gungor and Onur Gungor [7] was examined in detail. A meeting with Onur Gungor was held in Bogazici University to talk about his experiences on the issue. Detailed information about the game was discussed. The method and importance of reward system was analyzed [6]. Following are the some of the alternatives that were considered:

- Given a paragraph, determining how the paragraph is perceived by the user and drawing an inference from the relation of the certain structures and words within a paragraph with the perception of that paragraph by the user. The possible perceptions would include; formal, polite, cozy, slang, rude, funny, instructive, etc... To examine the results, a statistical analysis of the frequency of structures and words versus the perception of the paragraph might be useful.
- Given a paragraph containing words and idioms that is specific to certain region, determining which region or part of the country that excerpt would be from. Sound record is also an option that could be used instead of the paragraphs. The data that was collected during my previous studies would be used as an initial base for the study [10].
- Evaluating the words in Turkish in terms of their cultural value, meaning and how original it is. The collected data would be used in classification of the Turkish words in terms of these values.
- Given quotes from important people in history, evaluating the emotional effect of the quote on people. For instance, "That which doesn't kill us makes us stronger" Nietzsche, could be asked and one of the following possible answers could be received from the user; optimism, stability, endurance to changes in external circumstances. The results would be used to create a database of quotes classified as their emotional effect.
- Given a sentence that contains a passive verb, doing interpretation of implicit argument in passives. In some cases implicit argument would be an institution and in others it would be a specific person.
- Given a sentence that contains a verb in passive structure, determining whether it is actually a passive verb or a reflexive verb. The results would be used to determine in which sentences the verbs are used as passive, but an extensive information about the sentence that contain the verb is necessary.
- Given a sentence with a passive verb, interpretation o the implicit argument in the cases that the verb is an unaccusative passive verb or an aorist verb.

- Given a sentence with a unergative passive verb, interpretation o the implicit argument in the cases that the verb is a passive verb or an aorist verb.
- Given an intransitive Turkish verb, determining whether its passivized version could be used in the daily language.

Among the above alternatives, the last one is selected as the area of study. The reasons behind this selection were; it is a problem that can be solved easily by regular people but not by a regular computer, it is a unique problem that lets the creation of an enjoyable game to collect data from humans. Finally, no similar study on this issue was found on the literature.

3. PRE-STUDY AND GAME DESIGN

Before starting the design of the game, existing games that utilize from human computation were deeply analyzed. There were various games that were designed by Luis von Ahn. ESP Game [3] is one of the well-known games that he designed. This is a multiplayer game that matches random players from all over the world and asks them to guess the same image. There is a pre-defined time limit so that the game ends at a certain point. This limit also forces players to make a guess as fast as possible. So that the game could successfully get the first thought that the user has in his/her mind about the current image. As a result of the game, random images on the internet were labeled in a fast, accurate and free way. This popular game was a breakthrough in using human computation via games.

Verbosity [1] was another game designed by von Ahn. This was a game that again matches two players on the Internet and lets one of them give clues to the other player to guess a certain word. As a result of the game, detailed information about the properties of the object was collected. Another game of Luis von Ahn was Peekaboom [2], which let understanding the characteristic region of an image with a certain label. Tagatune [4] was another game of von Ahn that lets users tag a musical clip in a game. Phetch [5] was also designed by von Ahn. It let collecting descriptive explanations of images. Apart from the mentioned games, 3D games such as Second Life [12] were examined to obtained ideas in game design phase.

3.1. Comparison of the Game with Other Games in Literature

Given an intransitive Turkish verb, the game has the purpose of determining whether its passivized version could be used in the daily language. In this sense, it is unique in the literature. The answers that the game will come up will be useful for classification of intransitive verbs in terms of passivization capability and for further studies in the area. Additionally, on the contrary to the aforementioned games, *Dil Cambazi* is not a browser based game. This is a considerable advantage for people who do not have persistent internet connection, but want to play the game. The results can be submitted anytime by the player, so that the player can submit the results whenever s/he has a working internet connection. The game could be played in a wide range of platform since it's a java application.

Another advantage of the game is that, once registered, it does not require players to use the keyboard during the game. Players can play the game by using a mouse. Since it is relatively more cumbersome to use keyboard, such an approach was followed in the design.

3.2. Result Submission System and Methods of Submission

Result submission is the crucial part of the game. Results contain the data that we want to collect from the players. In order to keep the players playing the game, result submission should be practical. There are 3 ways to submit the result of the game.

The first method is to submit the results via clicking "Hemen Gönder" button at the end of the game. This process requires an active internet connection. In the case that the results are successfully received, the user is made aware of that by a short message. If the transmission fails, user is notified of this failure too.

If the first approach fails or if the user prefers not to send the results using previous option, s/he can use "E-posta ile Gönder" button to submit the results. In the case of this

option, the default e-mail client is activated and the <u>mailto</u>: part of the client is filled automatically with the e-mail address that we provided.

The last option is to send the result manually. In this option, user uses his/her preferred e-mail service to send an e-mail to us with the results attached to the e-mail. This manual option could be performed anytime.

The results are collected in a special Gmail account that was created for the use of the game. Later, these attachments are fetched from the e-mail server and collected in a local machine. After that the result files are combined in a single file and game results are extracted from here to create a statistical examination.

3.3. Time Limit of the Game

Time limit is an important part of the game. Most similar games with a purpose, uses a time limit in order to force the players answer within a certain time. It is a useful feature in two ways; first, it helps collecting the data which is the first thought of the users, which makes it a possible first thought of most other people too and secondly, it makes the game more challenging for the players so a user will want to play the game again. And finally, the point system will force the player to submit answers as correct as possible in a short time.

The time limit was determined as 240 seconds for 30 questions that were asked in both pre-game base data collection phase and in game phase.

4. COLLECTING BASE DATA FOR THE GAME

For the game to be played with a point and badge award system, a base result data is needed. Initially the game will be distributed to a number of anonymous users to get a base data for creating a robust award system.

4.1. Screenshots from the Initial Base Game

Note that following screenshots were taken from the initial version of the game. In this version the data was collected anonymously from a selected group of people. The design is plainer and contains fewer features. The results of the base data collection phase are analyzed later in the document.

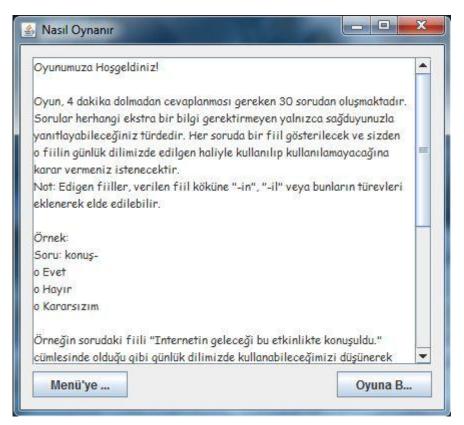
Welcome: It is the first screen that the game starts with. It has 3 buttons and a nonresizable window. Player can select appropriate button to continue. The buttons are self-explanatory. Player can return to "Welcome" screen if s/he chooses "Nasıl Oynanır" or "Hakkında" buttons. However, if the user clicks "Oyuna Başla", the game starts and player cannot return back to this screen.

👍 Hoşgeldir	iiz 🕒	. 🗆 🗙
Oyun	umuza Hoşgela	liniz!
	NASIL OYNANIR?	
	OYUNA BAŞLA	
	HAKKINDA	

About: The user can learn about the purpose of the game. It also contains useful information about data submission process.

Bu oyun Boğaziçi Üniversitesi	
akademik bir araştırmanın ürü	
Oyunu tamamladığınızda "HEA tıklayanak ya da "E DOSTA ile	GÖNDER" butonuna tiklayıp
oyun klasörünüzün altında oluş	
yollayarak siz de bu araştırma	TA SE THE 15 STREET BOOK
Yukarı <mark>d</mark> aki metodların boşarıs	17 almasi durumunda avun
sonucunda oluşturulacak "beni	
yollayabileceğiniz e-posta adr	
oyunboqazici@gmail.com	
oyunbogaziereginan.com	
Teşekkürler	
Menü'ye Dön	Oyuna Başla

How to Play: User can obtain detailed information about how to play the game in this instructive screen. It is possible to return to the welcome screen or start the game at this screen.



Question: A typical question that the user is expected to answer. The structure is straightforward. There is a question number, a verb as the question, a set of radio buttons for user to answer and a timer showing the remaining time. The color of the timer is blue to show that the user currently has plenty of time. User can go to the next question by answering the current question and clicking to "Sonraki" button.

🛃 Soru 🗖 🗖 💌 🗶
Soru: 11 sol-
 Evet Hayır Kararsızım
Sonraki Kalan Süre: 209

Question (less than 10 seconds left): in the case that the player could not complete the questions until the last 10 seconds, the timer turns to red to let the user be aware that the time is going to be over soon. If the user cannot complete the game in time, his/her answered questions are stored and the game goes to the final screen for submission of the result.

約 Soru		
Soru: 11	sol-	
) Evet		
Hayır		
🔘 Kararsızım		
Sonraki		Kalan Süre: 6

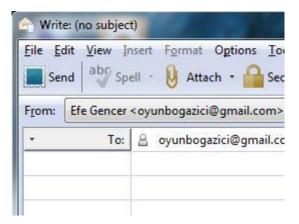
Finish the Game (or time is over): When the game is over, following screen with previously mentioned options is shown to the user. User can choose any option that s/he deems as appropriate.

Oyunumuzu Oynadığınız İçin Teşekkür Ederiz! Lütfen Sonuçları Bize Göndermeyi Unutmayın HEMEN GÖNDER E-POSTA ile GÖNDER ÇIKIŞ	🛓 Teşekkürler		_ _ ×
	Oyunun	nuzu Oynadığınız İçin Teşekkür E	deriz!
HEMEN GÖNDER E-POSTA ile GÖNDER ÇIKIŞ	Lütfe	n Sonuçları Bize Göndermeyi Unutma	ayın
REMEN GONDER E-POSTA ILE GONDER ÇIRIŞ	HEMEN CÔNDER		CIVIE
	HEMEN GÖNDER	E-POSTA ile GÖNDER	ÇIKIŞ

Results are sent: if the results are sent using "Hemen Gönder" option, the selections to send are deactivated and a message to inform the player is presented.

💁 Teşekkürler		— — X
	<mark>Sonucunuz Başarıyla Alınmıştır.</mark> Teşekkürler!	
HEMEN GÖNDER	E-POSTA ile GÖNDER	ÇIKIŞ

Send E-mail: If "E-posta ile Gönder" button is hit, then the default mail client is run in the system, ready to send the mail after attaching the result file.



4.2. Interpretation of the Collected Data from Anonymous Participants

Collected base data were combined in a local machine and the statistical results were generated at this machine. The answers of the participants were analyzed and some of them were selected as seeds. In selection process, common sense of the majority is used. These results are utilized in the award system.

4.3. Using Collected Base Data in Game Award System

In order to create a good award system, a set of results are needed to make a comparison with the submitted game results. These base results are used as "seeds" that will be asked to the players during the game without the participant noticing it. According to the user's answers to the seed questions, the weight of probable correctness of his/her other answers will be determined. Also a badge and point system were created. The results of the badges and points of the users are uploaded to a web site at every certain day of the week. So that users will have chance to learn their points and badges in weekly periods.

There are two aspects in the point and badge system. The first aspect is the points that changes from user to user at each game play. A better player will have higher chance to earn more point than a poor player at once. But in the long run, accumulated points of a user might be higher even though s/he is a worse player. On the other hand, badge system only awards the good player but not the poor player by checking user's correct answer rate to the seed questions.

The seed questions are used as factors in user success ratio calculation. There are five seed questions in each set of game-play period. Each of these questions has their specific weight calculated by ratio of most likely answer in collected base results. For instance a question that was answered 13 times "Evet", but only 1 time "Hayir" will have a weight of 0.9285 by scaling the rate to 1.

4.3.1. Badge system

Badges are specific named virtual prizes that a player may earn according to his/her actions in the game. Although there are some badges that are related to the points of players, this award system is mostly independent from the point system. User will know what to do to earn some of the badges and will not be aware of what s/he should exactly do to earn some other badges. This approach will be followed in order to make the game more mysterious and addictive.

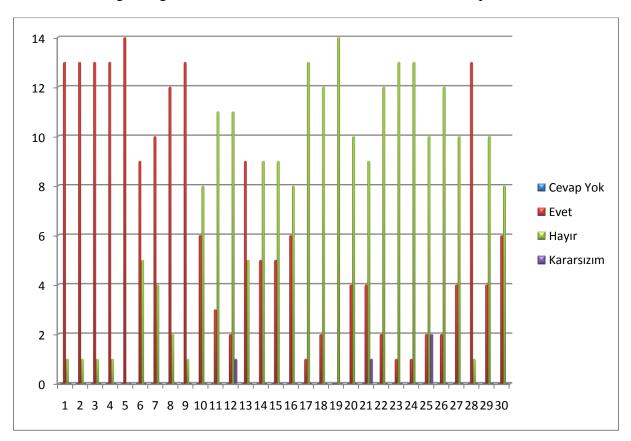
For instance one of the badges is "Bilge". It is one of the hardest badges that require a success ratio of 0.9 and a total point of 10000 or more.

4.3.2. Point system

Point system is a regular user award system. If the user plays a lot and at the same time answers the seed questions correctly, s/he has the chance to increase his/her points faster than others. It is another method of creating an addictive game. The collected results have shown that people are not tending to answer in a random careless manner. Most of the results were confirmed to be reasonably answered by checking the answers of the seed questions.

5. BASE DATA RESULTS

Base data collection period continued for 2 weeks and throughout the period, multiple participants played the game and submitted results via game data submission system. Total of 14 selected players had played the game and submitted their results.



Following histogram shows the distribution of the answers to 30 questions:

Following questions were asked in the base data collection process (in the order as in the histogram in figure):

Content of the Seed Questions

01 çalışmak> çalışıl-;	11 hıçkırmak> hıçkırıl-;	21 yaşlanmak> yaşlanıl-;
02 düşünmek> düşünül-;	12 horlamak> horlan-;	22 bunamak> bunan-;
03 koşmak> koşul-;	13 kızarmak> kızarıl-;	23 batmak> batıl-;
04 atlamak> atlan-;	14 öksürmek> öksürül-;	24 çürümek> çürün-;
05 oynamak> oynan-;	15 uyumak> uyun-;	25 donmak> donul-;
06 yürümek> yürün-;	16 ölmek> ölün-;	26 erimek> erin-;
07 yüzmek> yüzül-;	17 boğulmak> boğulun-;	27 kararmak> kararıl-;
08 ağlamak> ağlan-;	18 bayılmak> bayılın-;	28 kırmak> kırıl-;
09 gülmek> gülün-;	19 doğmak> doğun-;	29 patlamak> patlan-;
10 hapşırmak> hapşırıl-;	20 büyümek> büyün-;	30 solmak> solun-;

The results show that in most of the questions people tend to answer in the same way with other people. 22 of 30 questions were answered either as "Evet" or "Hayır" with a rate above 71.43%. This tendency to answer in a similar way is a useful result which lets us use these questions as "seed" questions in the game. They are used in the process of controlling whether players' answer the questions in a reasonable way. They are also used in determination of user's point and whether s/he earned a badge or not.

The rest of the questions (8 among 30) were not statistically good enough to be used in the process of judging the player's answer to decide whether s/he answers those questions correctly.

6. GAME PLAYING PERIOD

Game playing period was started after the completion of base data collection, interpretation of the base data results and completion of the new award system design. It will continue until enough data for making a correct, unbiased inference is possible.

Following sections investigates the details of the game.

6.1. Collecting User Submission with Known Users

On the contrary to the base data collection phase, the results shall be collected from *known* users. In this way, each user of the game will have chance to submit their result and learn their certain badges and weekly rank within the game community. For this purpose, registration system was added to the game.

In the first run of the game in a new system, registration system runs automatically to register the player to the game. Registration process requires a username chosen by the player. This username selection process creates a license file for the current user with registration username and 2 special user license IDs. These IDs are unique to each user and one of them is public, the other one is private. The public ID is used to let separation of distinct users with the possible same username. It is declared on the weekly result announcement together with the username. Following is the format of the weekly announcements, taken from 1st week results:

SIRA	ID	KULLANICI ADI	PUAN	BADGE
1	1449345071	xyz123	1000	Beyaz Tavşan ve Kaşif
2	870149297	s_duyar	800	Kaşif
3	571541565	immortalmad	500	Kaşif

HAFTANIN ENLERİ

6.2. Weekly Ranking and Badge Declarations

On weekly periods, the collected game results will be analyzed and new data will be added to our database. At a specified day of the week, the ranking of players within the game community and badges of the players will be declared. Users will have chance to learn the results and, hopefully, will play the game more to get a higher point or to earn a new badge upon learning their results.

The weekly result announcements are made in a web site. This web site is also used as a means of communication between players and us. The game could also be downloaded directly from this web site. It is a one-click download procedure with explanations. There is also a link to game's Twitter account in which users can follow game updates and latest news using Twitter.

The renewed design of the game contains more colorful appearance and aforementioned extra features.

6.3. Screenshots from the Game

Registration: If the user runs the game in a new machine that he did not run the downloaded game before, than a registration screen opens up and asks the user to fill the username section to be able to start the game.



Welcome: If the player is already registered, it is the first screen that the game starts with. Otherwise it is the second screen that appears after registration. It has 3 buttons and a non-resizable window. Player can select appropriate button to continue. The buttons are self-explanatory.



About: The user can learn about the purpose of the game. There are also useful information about data submission process.

Hakkında	x
Oyunumuza hoş geldiniz!	•
Bu oyun Boğaziçi Üniversitesi tarafından yürütülen akademik bir araştırmanın ürünüdür.	
Oyunun sonunda HEMEN GÖNDER butonuna tıklayarak ya da E- POSTA ile GÖNDER butonuna tıklayıp oyun klasörünüzün altında oluşturulacak dosyayı bize göndermeyi unutmayın.	I
Yukarıdaki metodların başarısız olması durumunda oyun sonucunda oluşturulacak "beniYolla" isimli dosyayı yollayabileceğiniz e-posta adresimiz:	
oyunbogazici@gmail.com	
İvi Fălenceleri	-
Menü'ye Dön Oyuna Başla	

How to Play: User can obtain detailed information about how to play the game in this instructive screen. It is possible to return to the welcome screen or start the game right there.

約 Nasil Oynanir	x
Oyunumuza hoş geldiniz!	
Bu oyun Türkçe bilginizle oynayabileceğiniz ve zamana karşı yarıştığınız bir oyun.	=
Tek yapmanız gereken ekrana yansıtılacak olan eylemin ve Türkçede alabileceği biçiminin kabul edilebilir olup olmadığına karar vermek.	
Ekranda yansıtılan eylem ve alabileceği biçimi Türkçede kabul edilebilir buluyorsanız EVET, kabul edilemez buluyorsanız HAYIR, karar veremediyseniz de KARARSIZIM seçeneklerine tıklamak bir başka kelimeye geçmek için yeterli olacaktır. 4 dakika dolmadan 30	
soruya yanıt vermelisiniz.	
Menü'ye Dön Oyuna Başl	a

Question: A typical question that the user is expected to answer. The structure is straightforward. There is a question number, a verb and its passivized version as the question, a set of radio buttons for user to answer and a timer showing the remaining time. The color of the timer is blue to show that the user currently has plenty of time. User can go to the next question by answering the current question and clicking to "Sonraki" button.

🐇 Soru	X
Soru 1: parçalan	nak> parçalan-
• 😇 Evet	
ି 🐑 _{Hayır}	
ି 🧐 Kararsızım	
Sonraki	Kalan Süre: 234

Question (less than 10 seconds left): in the case that the player could not complete the questions until the last 10 seconds, the timer turns to red to let the user be aware that the time is going to be over soon. If the user cannot complete the game in time, his/her answered questions are stored and the game goes to the final screen for submission of the result.

🛓 Soru 📒 🗖 🗙	
Soru 1: kocamak> kocan-	
• 📴 Evet	
ି 🐑 Hayır	
ି ତିହ୍ରି Kararsızım	
Sonraki Kalan Süre: 5	

Finish the Game (or time is over): When the game is over, following screen with previously mentioned options is shown to the user. User can choose any option that s/he deems as appropriate.

Tebrikler	efegencer!	
Lütfen Sonuçları G	öndermeyi Unutmayın	
HEMEN GÖNDER	E-POSTA ile GÖNDER	
HAFTANIN ENLERÍ	CIKIŞ	

Results are sent: if the results are sent using "Hemen Gönder" option, the selections to send are deactivated and a message to inform the player is presented.

💪 Teşekkürler		and the second second		×
	Sonucunuz Başa	rıyla Alınmıştır.		
	Teşek	kürler!		
[HEMEN GÖNDER	E-POSTA ile GÖNDER]	
	HAFTANIN ENLERI	ÇIKIŞ]	

Send E-mail: If "E-posta ile Gönder" button is hit, then the default mail client is run in the system, ready to send the mail after attaching the result file. (Content and subject fields are filled to remind of the need for adding the attachment)

and an or building a	nsert F <u>o</u> rmat Options <u>To</u> ell • 🕖 Attach • 🔒 Sec
From: Efe Gencer	<oyunbogazici@gmail.com></oyunbogazici@gmail.com>
• To:	a oyunbogazici@gmail.cc

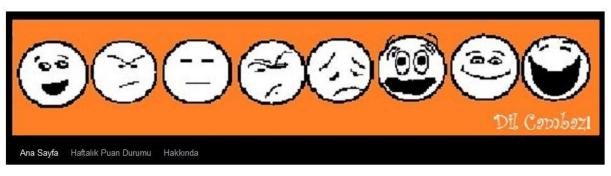
Haftanın Enleri: This button runs the browser and opens the weekly result announcement page. In the game window, user id is presented. User can check his/her username and match it with the user id and can see his/her result on the web site.

	ID: 675	5146717	
	Teşek	kürler!	
HE	MEN GÖNDER	E-POSTA ile GÖNDER]
HAI	FTANIN ENLERI	ÇIKIŞ	1

Game Blog: Weekly results and downloadable game will reside in the official blog of the game. Following is a snapshot from the official blog of the game at http://dilcambazi.wordpress.com/

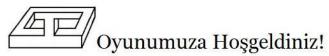
Dil Cambazı

Dil Cambazı Resmi Web Sitesi



Dil Cambazı'na Hoşgeldiniz!

Posted on 16 Mayıs 2011 by Dil Cambazı



Bu oyun Türkçe bilginizle oynayabileceğiniz ve zamana karşı yarıştığınız bir oyun. Oyunumuzu yan taraftaki Dil Cambazi v2.3.zip dosyasına tıklayarak indirebilirsiniz (Lütfen oyunun inmesi tamamlanana kadar bekleyiniz). İndirdiğiniz dosyayı zip içerisinden çıkararak oyun klasörüne ulaşabilirsiniz. Oyunu başlatmakta sorun yaşıyorsanız, lütfen kullandığınız java sürümünün güncel olup olmadığını <u>buradan</u> kontrol ediniz.

Tek yapmanız gereken ekrana yansıtılacak olan eylemin ve Türkçe'de alabileceği biçiminin kabul edilebilir olup olmadığına karar vermek.



7. CONCLUSION AND FUTURE WORK

NLP is a very active research area which is getting more and more popular. In contrast to traditional NLP applications, using human-computation in NLP applications leads to a more accurate and easy way of performing natural language processing. Collecting data through game playing is a satisfactory exchange for both parties. We, as the party that collects the data are satisfied because we collected a precious data through a way that would be extremely hard -or maybe impossible- with classical algorithmic methods. The other party, which is composed of game players, is also satisfied because they are just playing a game and competing with others to have fun.

Using the collected data in the base phase, the game design was changed and new features were added. A new award and point system were designed. A user registration system was created so that user recognition was successfully performed. It is considered to extend registration system by asking for information about the player (i.e. occupation, level of education)

REFERENCES

- von Ahn, L., Kedia, M., Blum, M.: Verbosity: a game for collecting common-sense facts. In: CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, New York, NY, USA, ACM (2006) 75-78
- von Ahn, L., Kedia, M., Blum, M.: Peekaboom: a game for locating objects in images. In: CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, New York, NY, USA, ACM (2006) 55-64
- von Ahn, L. Dabbish, L.: Labeling images with a computer game. In: CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM (2004) 319-326
- Law, E.L.M., von Ahn, L., Dannenberg, R.B., Crawford, M.: Tagatune: A game for music and sound annotation. In: ISMIR 2007: 8th International Conference on Music Information Retrieval. (2007)
- 5. von Ahn, L., Ginosar, S., Kedia, M., Liu, R., Blum, M.: Improving accessibility of the web with a computer game. In CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems, New York, NY, USA, ACM (2006) 79-82
- Choi, B., Lee, I., Choi, D., Kim, J.: Collaborate and Share: An Experimental Study of the Effects of Task and Reward Interdependencies in Online Games. CyberPsyChology & Behaviour. Volume: 10. (2007) 591-595
- Güngör, O., Güngör, T.: Morphological Annotation of a Corpus with a Collaborative Multiplayer Game. 11th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2010), Ed. A. Gelbukh, March 2010 Iaşi, Romania – LNCS (Lecture Notes in Computer Science), Vol.6008, 2010, p.74-85, Springer-Verlag, Berlin Heidelberg.
- 8. von Ahn, L.: "Human Computation." Wired Science (PBS, 2007), www.pbs.org/kcet/wiredscience/story/96-luis_von_ahn_human_computation.html.
- von Ahn, L.: "Human Computation" Tech Talk at Google in Mountain View, CA, USA, July 2006. <u>http://video.google.com/videoplay?docid=-</u> <u>8246463980976635143&q=human+computation#</u>
- 10. Gencer, A.E.: "Korkuteli Ağzı, Manileri, Türküleri ve Masalları". Kabataş Erkek Lisesi. İstanbul(2004).
- 11. Nakipoğlu-Demiralp, M. 2001. The referential properties of the implicit arguments of impersonal passive constructions. In Erguvanlı-Taylan, ed., 129-150.
- 12. Second Life, 2011. URL <u>http://secondlife.com/</u>.