Making NLP games with a purpose fun to play using Free to Play mechanics: RoboCorp case study

Dagmara Dziedzic
Department of Logic
and Cognitive Science
Adam Mickiewicz University
Wieniawskiego 1, Poznań, Poland
dagmara.dziedzic@amu.edu.pl

Wojciech Włodarczyk
Faculty of Mathematics
and Computer Science
Adam Mickiewicz University
Umultowska 87, Poznań, Poland
wojciech.wlodarczyk@amu.edu.pl

Abstract

The main aim of this talk is to present the new approach of employing the Free to Play (F2P) mechanics to games with a purpose (GWAPs) in order to provide a better playing experience for users. We have implemented and tested this solution in RoboCorp game. During the period from March 22 to May 25, 2015, 232 players tagged 3923 sentences in the game. The main objective in the game was annotation of the named entities retrieved from the Polish National Corpus. Gathered results proved this approach can be successfully used in NLP GWAPs.

1 Introduction

Games with a purpose (GWAP) are designed to meet two (distant at first sight) aims: a) to provide a player with an enjoyable game, and b) to serve as a tool to solve a certain difficult problem. Since the term coinage and the first successful GWAP implementation by Von Ahn and Dabbish (2008), we have been witnessing a growing popularity of game-based approaches to the widely understood problem solving in many different domains including science.

2 Related works

There are GWAPs that were used to solve different NLP problems. They have been used, for example, to gamify the task of anaphora resolution (Poesio et al., 2013) and the wordsense disambiguation (Venhuizen et al., 2013). However, as pointed out by Jurgens and Navigli (2014), most of them incorporate only popular and simple elements of the game mechanics such as points, achievements, rankings, and levels. GWAPs that are used in the NLP domain are focused on text; thus, they strongly resemble traditional tools used by linguists to perform annotation (Jurgens and Navigli, 2014).

One of the possible solutions that can be used to provide a more engaging gameplay is to replace the traditional text annotation with pictures complemented with an dynamic, arcade elements (for more see Puzzle Racer and Ka-boom! by Jurgens and Navigli (2014)). Although games of this type can be more attractive for players, replacing the text annotation with arcade elements may lead to a situation where the quality of results will depend on the players reflexes, not their linguistic competence and thoughtful decisions.

3 RoboCorp

The aforementioned problems were one of the main motivations for developing the new approach, which allows us to introduce a fun, gameplay mechanics to GWAPs. We have implemented it and tested in a project called RoboCorp[1]. The main purpose of this project was to check whether the players led only by their intuition and language competence could determine named entities in the Polish sentences at least as well as language experts. The main objective in the game was annotation of the named entities retrieved from the Polish National Corpus. In order to bring the playfulness to the player, certain mechanisms known from Free2Play (hereafter F2P) games were implemented in RoboCorp.

F2P games offer free of charge fun for players. The idea is that they can use additional features of the game design in order to encourage play-

[1]The project was developed in the Unity environment and it is available in versions for browsers (http://citygames.zetsystem.com.pl/gwap/GWAP/) and mobile devices with the Android operating system (https://play.google.com/store/apps/details?id=com.citygames.robocorp).

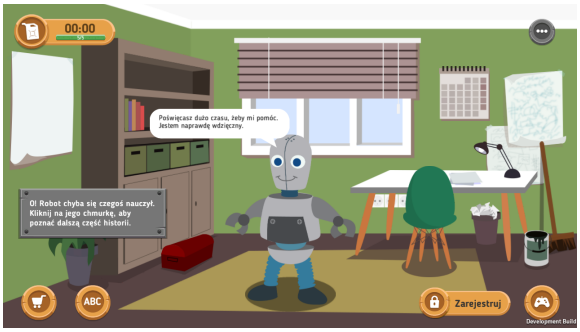

Figure 1: Main view of RoboCorp game in which the robot is trying to communicate with a player using encrypted messages

ers to purchase items in the game. The payments are small, which is why they are called micro-payments. F2P games implement several techniques which aim at improving the retention of players – the desired outcome is that players want to return to the game often.

The main character in the RoboCorp game is a little robot. There is also a plot for the game. The robot cannot speak Polish. In order to communicate with him and help him get back home, the player has to annotate sentences. The more named entities you will annotate, the more understandable the robot becomes. Annotation is also rewarded with points. To provide more fun and an opportunity to take a break from the annotation without leaving RoboCorp a mini-game was added – a platform game, where we try to send the robot home. In order to do this the player has to gain fuel and upgrades, which can be bought by points he/she earns by annotating sentences in appropriate module of the game. So in consequence, the currency in the game is not real money as in the case of classical F2P games, but your work (annotation). RoboCorp uses the idea which we will call micro-work (by analogy with micro-payments).

What is important, enriching the game with the mechanics known from F2P games does not affect the quality of the annotations. Game tests and the analysis of obtained data suggest that enriching a GWAP game with additional elements inspired with F2P genre makes the game more interesting without any loss in the quality of the gathered data (for detailed discussion see Dziedzic (2016)). 232 RoboCorp players annotated a set of 2.000 unique sentences.

Players annotated a total of 3923 annotations in the game. This means that each of the sentences has been marked by about two players (average of 1.96 per player, with a standard deviation of 0.99). At the end of project, there were 763 sentences that were annotated only one time (whereas some were annotated even seven times), which shows that the sampling algorithm should be enhanced.

83% (precision computed for all the categories) of proper names units were annotated in the same way as the languages experts did in the case of the golden standard. Additionally, players had annotated the 87% (recall computed for all the categories) of all proper names units.

Table 1: *The quality measure of the result corpus in comparison with the gold standard*

Named entities	Precision	Recall	*f-score*
Personal names	0.9515	0.8968	0.9234
Locations	0.9323	0.6768	0.7842
Names of organisations	0.7549	0.7603	0.7576
Temporal expressions	0.9883	0.7775	0.8703
Overall	0.9112	0.8772	0.8938

The proposed approach not only fulfills its role as an attractive research tool, but also provides an entertainment for the players in the form of an arcade game. With regard to ways of encouraging players to annotate the data, the project RoboCorp constitutes a novelty among other GWAPs in the NLP domain, but due to modular architecture of this approach it can be also easily implement in other domains.

References

D. Dziedzic. 2016. Use of the Free to Play model in games with a purpose: the RoboCorp game case study. *Bio-Algorithms and Med-Systems*, 12(4):187–197.

D. Jurgens and R. Navigli. 2014. Its all fun and games until someone annotates: Video games with a purpose for linguistic annotation. *Transactions of the Association for Computational Linguistics (TACL)*, 2:445–464.

M. Poesio, J. Chamberlain, U. Kruschwitz, L. Robaldo, and L. Ducceschi. 2013. Phrase Detectives: Utilizing collective intelligence for internet-scale language resource creation. *Transactions on Interactive Intelligent Systems (TIIS)*, 3(1):1–44.

N. Venhuizen, V. Basile, K. Evang, and J. Bos. 2013. Gamification for word sense labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers*, pages 397–403.

Luis Von Ahn and Laura Dabbish. 2008. Designing games with a purpose. *Communications of the ACM*, 51(8):58–67.