Modelling Word Associations and Interactiveness for Describer Agents in Word-Guessing Games

A Case Study for the Location Taboo Game

Verna Dankers
10761225

Bachelor Thesis Honours Extension
Credits: 6 EC

Bachelor Opleiding Kunstmatige Intelligentie
University of Amsterdam
Faculty of Science
Science Park 904
1098 XH Amsterdam

Supervisors
Dr. Ayseun Bilgin
Dr. Raquel Fernández

Institute for Language and Logic
Faculty of Science
University of Amsterdam
Science Park 904
1098 XH Amsterdam

July 30, 2017
Abstract

In word-guessing games, one player describes a stimulus and his partner should try to guess what the stimulus is. Producing a response to such a stimulus requires an associative mechanism. Additionally, the production of associations that allow the partner to guess what the stimulus is requires modelling a shared context. The Location Taboo Game is such a word-guessing game, in which a describer gives simple textual hints about a target city, and the other player should try to guess this city. The hints given should not contain words from a list of taboo words. In this thesis, an architecture for an artificial describer agent is presented and evaluated through simulation and an empirical study. To be able to elicit a correct guess from a human guesser, the artificial describer agent should mimic associations that humans have with geographical locations. The artificial describer agent extracts word associations from a semantic vector space that has been created with a context-predicting distributional semantic model. Firstly, two methods for extracting word associations are detailed: a nearest neighbours approach that uses the list of taboo words, and an approach that applies clustering and analogical reasoning on a dataset of games played by humans. Secondly, interactiveness is modelled through a rule-based approach for the generation of clues that depend upon the guesser’s response. Different variants of the methods for clue generation are evaluated through simulation and an empirical study. These describer agents could elicit a correct guess for 37.86% of the games and for 50.00% of the games for which the human player knew the target city.
Contents

1 Introduction 4

2 Background 5
   2.1 Theoretical Foundation ........................................ 5
   2.2 Related Work on Word-Guessing Games ........................ 5

3 Case Study: Location Taboo Game 6

4 Method 6
   4.1 Describer Agent .................................................... 7
   4.2 Creation of Independent Clues ..................................... 7
   4.3 Creation of Dependent Clues ..................................... 10
   4.4 Implementation .................................................... 11

5 Experiments and Results 13
   5.1 Automatic Evaluation ............................................ 13
   5.2 Empirical Evaluation ............................................. 15

6 Discussion 17

7 Conclusion and Future Work 18

References 19

Appendix A Questionnaire 20
   A.1 Opening Questionnaire ........................................... 20
   A.2 Questions per Game ............................................... 20
   A.3 Closing Questionnaire .......................................... 21

Appendix B Empirical Evaluation 22
   B.1 Results per Game ................................................. 22
   B.2 Results per Participant ......................................... 23
1 Introduction

When a person is asked to give an associative response to a stimulus, he goes through three stages: understanding the stimulus, operating upon it according to its meaning and producing an answer. Within the second stage of operating upon the stimulus, an associative mechanism is used to determine the response (Clark, 1970). When the association game is played by multiple people, who should guess the target, the associative mechanism should not only produce a response to the stimulus, but also take into account the players’ shared context. The analysis of word association games reveals properties of linguistic mechanisms underlying the associative mechanism and can contribute to the creation of tools for learning new languages to artificial agents (Steels, 2001).

One of the games in which associative mechanisms play a major role, is the Taboo word-guessing game, in which a describer agent provides clues about the term to be guessed without mentioning the target term or any of the related terms from a list of taboo words. Thus, the game requires the describer to think of well-known facts about the target word to enable the guesser to guess it correctly. The Location Taboo Game (LTG) is a particular version of Taboo in which the target terms are cities and clues are simple and short textual hints. The LTG has been invented for the Taboo City Challenge,¹ which is a competition inspired by Taboo, where artificial guesser agents play the LTG.

This work builds upon the thesis of Dankers (2017), in which an architecture for an artificial guesser agent (AGA) for the LTG has been proposed. Here, an architecture for the artificial describer agent (ADA) is presented. Additionally, a rule-based approach for interactive game playing behaviour is suggested. The describer agent uses a vector space, created with the context-predicting distributional semantic models (DSM) named Skip-Gram, to model knowledge about cities. Three different methods of extracting city descriptions from this vector space are proposed. These different methods are evaluated through simulation and through an empirical study. The describer’s objective is to elicit a correct guess from its users. Thus, the ADA should be able to produce words associated with the target location that are recognisable for the guesser agent.

Through the creation of an ADA that generates clues for target cities in the LTG and the automatic and empirical evaluation of the architecture, the following research question is investigated:

How can word vector spaces facilitate generating human-like word associations for a Describer Agent in the Location Taboo Game?

The creation and evaluation of the architecture can be broken down into different components, which leads to the following subquestions:

1 To what extent can the artificial describer agent generate descriptions for the Location Taboo Game that are recognisable for humans?
2 How can word vector spaces be used to extract syntagmatic and paradigmatic word associations for cities?
3 How can the Location Taboo Game be made interactive for both the describer and the guesser agent?

¹The challenge is organised by the ESSENCE Network https://www.essence-network.com/challenge/.
The outline of this thesis is as follows: In the following section, a theoretical foundation is provided and related work is reviewed. Section 3 discusses the rules of the LTG. Section 4 presents the ADA architecture and describes the different methods for clue generation. An evaluation of the architecture is given in Section 5. The results are discussed in Section 6. Some concluding remarks and suggestions for future research are given in Section 7.

2 Background

2.1 Theoretical Foundation

The functional relations between words are traditionally divided into two main kinds: syntagmatic and paradigmatic relations (De Saussure, 1916). Syntagmatic relations concern positioning, and relate entities that co-occur in the text. This relation is a linear one, and applies to linguistic entities that occur in sequential combinations. Syntagmatic relations are combinatorial relations, which means that words that enter into such relations can be combined with each other and often fall into different syntactic categories, such as ‘plant’ and ‘green’. Paradigmatic relations, on the other hand, concern substitution, and relate entities that do not co-occur in the text, but that do occur in the same context (Sahlgren, 2006), such as ‘plant’ and ‘flower’. Words with paradigmatic relations often fall in the same syntactic category. Thus, words have a syntagmatic relation if they co-occur, and a paradigmatic relation if they share neighbors.

These relations play an important role in the free-association or word-association game, in which a stimulus is presented to a subject that responds with the first thing that comes to mind. These kind of games reveal properties of linguistic mechanisms underlying it. The ability to produce word associations is derived from the ability to generate and understand natural language. For humans, paradigmatic associations are more common than syntagmatic associations (Clark, 1970).

DSMs represent words as real-valued vectors based upon the context in which the words appear (Baroni et al., 2014). Therefore, the vector space groups paradigms, which are sets of substitutable entities.

2.2 Related Work on Word-Guessing Games

Multiple clue givers have been created for word-guessing games before. One of those systems is Daboo, in which a describer agent is modelled for a Taboo game in which only nouns are used as target terms (Marti & Emmett, n.d.). The system can describe five nouns in total. The research mostly focused on creating an interactive game, rather than being able to describe a large amount of target terms. Daboo always immediately responds, starting with ‘ummmm’ to acknowledge the input while the evaluation algorithms are run. The system can indicate how close the user’s guess was through a phrase such as ‘You are very close...’ or move towards a different describing strategy through ‘Think of this...’. The user’s knowledge is modelled by matching his guess to one of six knowledge domains and using the accompanying describing strategy. The closeness of the guess to the taboo word is modelled through their semantic relationship, for which Daboo relies on WordNet.

Verbosity, a system that implements Taboo as well, has been proposed by Von Ahn et al. (2006) to test and collect common-sense facts. In Verbosity a narrator, the clue giver, and a guesser interact. The clues cannot contain the target term, and have to be given according to sentence templates such as ‘[...] is a kind of [...]’; in which the narrator can fill in the
blanks. The game has been played by pairs of humans on-line, and this data is used to create an automated narrator. This automated narrator collects a set of facts for a given target term from the dataset of games and presents a subset to the guesser. Based on the success of the game, the scores of the facts are increased or decreased. The score gives an indication of the quality of the clue regarding the target term. Of 200 randomly selected guesses, 85% was rated as accurate for the target term, upon manual analysis.

Heath et al. (2013) reviewed methods for building both describer and guesser agents for a word-guessing game named Wordlery, which is similar to Taboo. In their work, two methods for obtaining word associations were considered: using human free association norms and applying count-based semantic models to build word vectors from a corpus. The describer agent describes the hidden word by presenting the user with the top \( n \) word associations for the hidden word. For the count-based semantic models these are terms that are close to the hidden word in the vector space, which are terms with a paradigmatic relation to the hidden word. The models were evaluated by playing the game. The free association norms outperformed the count-based semantic models. Combining the two methods of forming word associations was superior to each of the methods in isolation. Heath et al. (2013) suggested that more advanced corpus-based semantic models, which take into account additional semantic information, may improve the results on similar tasks. This thesis builds upon this work by using a context-predicting type of DSM and by extracting both paradigmatic and syntagmatic associations from the vector space.

3 Case Study: Location Taboo Game

In this section, the rules for the LTG are laid out. A more detailed description of the game has been presented by Adrian et al. (2015). An LTG is played by a describer agent and a guesser agent. Hints are simple English noun phrases, consisting of one to three words that are common nouns, adjectives, or connectors. The hints may not include proper nouns. For example, if ‘Verona’ is the target city, the clue ‘Romeo and Juliet’ is not allowed, but ‘tragic love story’ is. Although there is no closed set of cities available for the LTG, it should only concern well-known cities.

The describer chooses a target location and starts the game by providing the first hint. Based on this hint, the guesser tries to guess the city that is being described. As long as the guess is incorrect, the describer provides a new hint and the game continues until there are no more hints left. If the guesser has not been able to find the right city before the describer runs out of hints, the game is considered to have failed. An example game is shown in Table 1.

Several dataset of games are made available by the ESSENCE Network. The first dataset is a set of 226 cities with their taboo words (dataset 1). Secondly, 82 real-world games are provided, with the taboo words, the hints provided by human describers and the guesses of human guessers (dataset 2). Thirdly, a set of 149 games was made available through an API\(^2\) for the Taboo City Challenge, containing hints provided by human describers (dataset 3). The corresponding taboo words are included in the first set of 226 cities.

4 Method

This section describes the describer agent’s architecture. Firstly, a general overview of the agent is presented. The describer agent can give the guesser agent two types of clues: clues

\(^2\)http://challenge.essence-network.com/describer/v1/
Table 1: An example LTG played by a human describer and a human guesser, for which the target city is Rio de Janeiro. The corresponding taboo words are ‘beaches’, ‘carnival’, ‘monuments’, ‘bananas’, ‘parade’, ‘crowds’, ‘dance’, ‘soccer’ and ‘poverty’.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describer</td>
<td>huge statue</td>
</tr>
<tr>
<td>Guesser</td>
<td>New York</td>
</tr>
<tr>
<td>Describer</td>
<td>festival</td>
</tr>
<tr>
<td>Guesser</td>
<td>Buenos Aires</td>
</tr>
<tr>
<td>Describer</td>
<td>animated movie</td>
</tr>
<tr>
<td>Guesser</td>
<td>Rio de Janeiro</td>
</tr>
</tbody>
</table>

dependent upon the user’s guess, or independent clues, describing general features of the target city. The different strategies employed for the generation of independent clues are presented in Subsection 4.2 and the rule-based approach for modelling interactiveness through the generation of dependent clues is given in Subsection 4.3.

4.1 Describer Agent

The describer agent prepares itself for a new game by choosing a target city, either in a static order from a list or randomly, and loads the clue vectors for the target city. During the game these independent clues will be extracted in the order of their similarity to the target city.

The describer agent starts the game by presenting the first hint to the guesser. Thereafter, the describer waits for an answer from the guesser. The answer should contain a guess (a location) as an answer to a hint presented by the describer. The game concludes successfully if the guess matches the target location. In this case, the describer informs the guesser of his success by mentioning ‘Correct’, selects a different target and repeats the process. If the location was not guessed correctly, the describer selects the next hint and waits for a guess. This process continues until there are no more hints to be presented to the guesser or until the maximum number of game iterations, $i$, has been reached. For the experiments conducted, $i$ has been set to 12. In this case, the describer informs the guesser of his failure by mentioning ‘Incorrect, the target city was [...]’; selects another target city and the game resumes.

4.2 Creation of Independent Clues

For every city, a set of clues with general facts about the target city can be prepared, independent of the guesser agent’s input. These facts are extracted from the word vector space and with these facts vector subspace is created for every target city. When playing the game, the ADA uses this subspace to extract clues very quickly. The rules of the LTG only allow for clues with a syntagmatic relation to the target city. As word vector spaces normally group terms with paradigmatic relations, one cannot simply gather nearest neighbours in the vector space. Three alternative approaches are suggested.

The vectors used were originally prepared for the architecture of the guesser agent developed by Dankers et al. (2017). These vectors were created from a corpus that consists of Wikipedia and Wikivoyage (Wiki*) pages related to cities and countries, with the Skip-Gram algorithm (Mikolov et al., 2013).
Using Taboo Words (TB) The first strategy for the generation of independent clues relies on the fact that the taboo words given in the game have very strong syntagmatic associations to the target city, as their exclusion serves to complicate describing the target city for the describer agent. Terms that are closely related to those taboo words are therefore most likely relevant for the target city as well. For every taboo word of the target city, the neighbour with the highest similarity to both the target city and the taboo word is selected as a clue, if that term is allowed according to the rules specified for the LTG. If there is no allowed term within the 100 nearest neighbours that taboo word is not used in the game. Example clues for the target city Paris, extracted using the provided taboo words from dataset 1, are presentend in Table 2.

The example illustrates that proximity in the vector space does not necessarily result in terms that describe the target city accurately—e.g., ‘second largest city’ is incorrect for Paris. A second disadvantage of this strategy is that it strongly relies on the availability of taboo words data for target cities. This restricts the target cities that can be described by the agent and restricts the number of clues that can be generated per city.

Table 2: Taboo words and their nearest allowed neighbours for the target city ‘Paris’, as found through the strategy of using taboo words within the Wiki* vector space.

<table>
<thead>
<tr>
<th>Taboo</th>
<th>Neighbour</th>
</tr>
</thead>
<tbody>
<tr>
<td>capital</td>
<td>second largest city</td>
</tr>
<tr>
<td>cafes</td>
<td>coffee shops</td>
</tr>
<tr>
<td>rivers</td>
<td>tributaries</td>
</tr>
<tr>
<td>museums</td>
<td>attractions</td>
</tr>
<tr>
<td>cathedrals</td>
<td>churches</td>
</tr>
<tr>
<td>art</td>
<td>galleries</td>
</tr>
<tr>
<td>monuments</td>
<td>landmarks</td>
</tr>
<tr>
<td>towers</td>
<td>skyscrapers</td>
</tr>
<tr>
<td>palaces</td>
<td>opulent</td>
</tr>
<tr>
<td>baguettes</td>
<td>three star restaurants</td>
</tr>
<tr>
<td>pastries</td>
<td>croissants</td>
</tr>
<tr>
<td>fashion</td>
<td>luxury boutiques</td>
</tr>
</tbody>
</table>

Using Taboo Words and AutoExtend (TBAE) The second strategy builds upon the first one, as it uses the neighbours of the taboo words and uses the nearest neighbours in a vector space of synonym sets in addition. Using just the neighbours from the regular vector space can result in multiple relations between the clue and the actual taboo word, since, for example, both antonyms and synonyms can lie in close proximity to each other in the word vector space. This method attempts to ensure that synonyms of taboo words are included in clues presented to the user.

For this, vectors created with the AutoExtend$^3$ (AE) algorithm are used that extend an existing word vector space with WordNet’s$^4$ lexemes and synsets (Rothe & Schütze, 2015). This extension method relies upon the so-called synset constraints: both words and synsets are sums of their lexemes. These constraints are used to represent the WordNet synsets and lexemes in the same vector space as a vector space inputted to the AE algorithm. Here, the

---

$^3$http://www.cis.lmu.de/~sascha/AutoExtend/

$^4$http://wordnetweb.princeton.edu/perl/webwn
Google News word2vec vectors\(^5\) and their extended space with synsets are used for this second strategy. A new clue is composed of two parts: the nearest allowed neighbour of the taboo word in the regular vector space, and the first term from the nearest allowed neighbour synset of the taboo word in the synset vector space. Using the extended vector space ensures that a word synonymous to the taboo word is included in the clue. An example of this strategy is shown in Table 3.

Table 3: Taboo words and their nearest allowed neighbours for the target city ‘Amsterdam’, as found through the strategy of using taboo words and the AE algorithm within the Google News vector space.

<table>
<thead>
<tr>
<th>Taboo</th>
<th>Neighbour</th>
<th>Neighbour from Synset Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>canals</td>
<td>waste slops</td>
<td>watercourse</td>
</tr>
<tr>
<td>art</td>
<td>postimpressionist</td>
<td>graphics</td>
</tr>
<tr>
<td>museums</td>
<td>aquariums zoos</td>
<td>depository</td>
</tr>
<tr>
<td>weed</td>
<td>cocaine</td>
<td>narcotic</td>
</tr>
<tr>
<td>prostitutes</td>
<td>streetwalkers</td>
<td>harlotry</td>
</tr>
<tr>
<td>wind mills</td>
<td>velvetleaf</td>
<td>nut sedge</td>
</tr>
<tr>
<td>tulips</td>
<td>daffodils hyacinths</td>
<td>flower</td>
</tr>
</tbody>
</table>

Using Clue Clustering (C-IDF, C-COS)  This last strategy for the generation of clues does not directly use taboo words prepared for the target city, but uses data from games played before to infer new associations for the current target city. The clues from games played by humans often fall into distinct categories such as historic events, local cuisine, environmental and geographical features and local places of interest. This strategy uses this notion to extract clues from different parts of the vector space, where clusters of words represent those categories.

Firstly, clues from 100 games from dataset 3 were set aside and terms that were uninformative without context, such as ‘larger’ or ‘small’, were removed. Secondly, the Wiki\(^*\) vectors of those clues were clustered into 10 groups with the k-means clustering algorithm.\(^6\) Thirdly, these clusters were used to create clues for new target cities, through analogical reasoning. For every clue within a cluster, the corresponding city was used in an analogy with the target city – e.g. for the target city Prague the analogy with Venice and its clue ‘pasta’ gives the clue ‘potato salad’. This resulted in many analogies per cluster. Two metrics are used to extract the optimal clue per cluster: the inverse document frequency (IDF) for the Wiki\(^*\) corpus and the Cosine similarity to the target city. These two metrics are used to create two different variants of this strategy: clustering and selecting through IDF (C-IDF) and clustering and selecting through the Cosine similarity (C-COS). The analogies are extracted from the vector space according to the Cosmul similarity (Equation 1), which is state-of-the-art in analogy recovery (Levy et al., 2014):

\[
\text{arg max}_{b \in V} \frac{\cos(b*, a*) \cos(b*, b)}{\cos(b*, a)}
\]

where the analogy is \(a\) is to \(b\) as \(a*\) is to \(b*\) and \(V\) represents the vocabulary from which the terms are extracted. The clusters were labelled with a general topic manually and the

\(^5\)The vectors are available at: https://code.google.com/archive/p/word2vec/.

clusters are visualised in Figure 1. An example description for the city *Prague* is presented in Table 4.

![Figure 1: The 2-dimensional PCA projection of 300-dimensional vectors from the different clusters and a small set of cities.](image)

Table 4: Clusters and their analogies found for the target city *Prague*, prepared with the clustering strategy with the Wiki* vector space.

<table>
<thead>
<tr>
<th>Cluster Label</th>
<th>Clue for IDF</th>
<th>Clue for Cosine similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buildings</td>
<td>best explored walking</td>
<td>best explored walking</td>
</tr>
<tr>
<td>Nature</td>
<td>occasionally snowy winters</td>
<td>highly unpredictable</td>
</tr>
<tr>
<td>Food</td>
<td>garlic soup</td>
<td>vegetarian meals</td>
</tr>
<tr>
<td>Sports</td>
<td>single ticket valid</td>
<td>single ticket valid</td>
</tr>
<tr>
<td>Environment</td>
<td>blue eyed</td>
<td>bad</td>
</tr>
<tr>
<td>(Alcoholic) Beverages</td>
<td>black tea</td>
<td>beers</td>
</tr>
<tr>
<td>Terror</td>
<td>thirty years war</td>
<td>thirty years war</td>
</tr>
<tr>
<td>Politics</td>
<td>protocol signed</td>
<td>military alliance</td>
</tr>
<tr>
<td>Media &amp; Culture</td>
<td>world men handball</td>
<td>world men handball</td>
</tr>
<tr>
<td>History</td>
<td>wollastonite</td>
<td>carpets</td>
</tr>
</tbody>
</table>

### 4.3 Creation of Dependent Clues

**Describer Agent** Within the dataset of 82 real-world games played by human describer and guesser agents, multiple hints are included that depend solely upon the guesses of the human guesser agent. However, they carry no value when evaluated without the data of the interactive game. To capture interactive game playing behaviour adequately, the artificial describer agent should be able to guide the guesser agent into the right direction, by adapting the clues given during the game.
Based upon a manual data analysis for the 82 real-world games (dataset 2) for which both clues and guesses are recorded, 23 clues were identified as containing only information dependent on the guesses provided by humans. This is 16% of the clues, the first hints of every game excluded. 19 of those clues capture a geographical aspect, either the location of a city (‘eastern’, ‘continent’ or ‘closer’) or its population size (‘smaller’ or ‘largest city’). The four remaining clues were ‘different accent’, ‘more obscure’, ‘older’ and ‘older’. Because the majority of the dependent clues captures geographical relations and these aspects are more easily generalisable than relations captured in the four remaining clues, the geographical aspect is used to create interactive game playing behaviour.

Geographical relations between cities are objectifiable in multiple manners. Here, the identification of borders, both continents and countries, and the geodesic distance are used as measures to compare different locations. To simplify the notion of dependent clues, we assume that a dependent clue is only based on the last guess given by the guesser agent. The following dependent clues can be given: If the city guessed is on the wrong continent, the dependent clue is ‘different continent’. If the city guessed lies on the right continent, but within the wrong country, the dependent clue becomes ‘different country’. If the city guessed is in the right country and lies within $\alpha$ kilometres of the target city, the dependent clue is ‘close’, otherwise the direction of the target city with regard to the last guess is indicated with its cardinal direction. For example, if the target city is Oxford and the user guesses Brighton, the dependent clue can be ‘northerwestern’. Whether the describer agent outputs an independent or a dependent clue, is based on a probability $\beta$, which has been set to 20% to approximate the percentage of dependent clues found through the manual data inspection.

**Guesser Agent** The AGA architecture created for the bachelor thesis has been extended by a mechanism for the interpretation of dependent clues. The AGA can interpret ‘different continent’, ‘different country’, ‘close’ and the cardinal directions – e.g. ‘southeastern’ or ‘northern’.

If the AGA receives the hint ‘different continent’, all countries that lie on the same continent as the city last guessed, are removed from the list of candidate cities. Similarly, for ‘different country’, all cities that lie in the same country as the city last guessed, are removed. For the clue ‘close’ the list of candidate cities is reduced to the cities that lie within a radius of $\alpha$ kilometres from the city last guessed. If the clue received is a cardinal direction, all cities that do not match that direction when compared to the city last guessed, are removed from the list of candidate cities.

### 4.4 Implementation

The ADA architecture is implemented in Python 2.7. The `gensim.models.word2vec` library is used to load the word vectors, to calculate nearest neighbours and to perform analogical reasoning. The `geonamescache` library is used to extract the geographical data needed to generate dependent clues. An overview of the main algorithm of the ADA when playing the LTG is presented in Algorithm 1.

The strategies for the generation of independent clues have been explained in Section 4.2. Algorithm 2 details the implementation of the strategies using taboo words. Algorithm 3 details the implementation of the clustering strategy. Clues are allowed if they conform to the LTG game rules, as explained in Section 3:

---

7https://radimrehurek.com/gensim/models/word2vec.html
8https://github.com/yaph/geonamescache
1. The clue does not include one of the taboo words or words directly referring to the target
term, which is implemented by stemming and lemmatisation with the PorterStemmer\(^9\)
and the WordNetLemmatizer\(^10\);

2. The clue is not the name of a country or city, which is implemented by matching clues
to the databases of the geonamescache and the geotext\(^11\) libraries;

3. The clue does not contain a pronoun, which is implemented through Part-Of-Speech
tagging with the StanfordPOSTagger\(^12\);

4. The clue only contains English words or words commonly used in English, which is
implemented by checking that the word is an English word according to the enchant\(^13\)
library or by checking that the word is present in the nltk.corpus\(^14\) corpus.

---

**Algorithm 1:** General Game Playing Algorithm

\[
\text{Input} : \text{target\_city, subspace\_target\_city, guess, } \alpha, \beta
\]

\[
\text{if } \text{iteration} == 1 \text{ or giveIndependentClue}(1-\beta) \text{ or not isCityName(guess) then}
\]
\[
\quad \text{clue} \leftarrow \text{mostSimilar(subspace\_target\_city)}
\]

\[
\text{else}
\]
\[
\quad \text{if } \text{continentCorrect(guess) then}
\]
\[
\quad \quad \text{if } \text{countryCorrect(guess) then}
\]
\[
\quad \quad \quad \text{distance} \leftarrow \text{getDistance(guess, target\_city)}
\]
\[
\quad \quad \quad \text{if } \text{distance} < \alpha \text{ then}
\]
\[
\quad \quad \quad \quad \text{clue} \leftarrow \text{‘close’}
\]
\[
\quad \quad \quad \text{else}
\]
\[
\quad \quad \quad \quad \quad \text{clue} \leftarrow \text{suggestCardinalDirection(guess, target\_city)}
\]
\[
\quad \quad \text{else}
\]
\[
\quad \quad \quad \text{clue} \leftarrow \text{‘different country’}
\]
\[
\text{end}
\]

\[
\text{else}
\]
\[
\quad \text{clue} \leftarrow \text{‘different continent’}
\]
\[
\text{end}
\]

\[
\text{return } \text{clue}
\]

---

\(^9\)http://www.nltk.org/_modules/nltk/stem/porter.html
\(^10\)http://www.nltk.org/_modules/nltk/stem/wordnet.html
\(^11\)https://github.com/elyase/geotext
\(^12\)http://www.nltk.org/_modules/nltk/tag/stanford.html
\(^13\)https://github.com/AbiWord/enchant
\(^14\)http://www.nltk.org/api/nltk.corpus.html
Algorithm 2: Using Taboo Words (and AE) Strategy

Input : vectors, target_city, taboo_words, autoextend, synset_vectors
clues ← emptyList()
for taboo in taboo_words do
    neighbours ← vectors.most.similar([taboo, target_city], 100)
    for term in neighbours do
        if allowed(term, taboo_words) then
            if autoextend then
                extension ← nearestAllowedNeighbour(synset_vectors, taboo, taboo_words)
            end
            term ← join(term, extension)
        end
        clues ← add(clues, term)
        break
    end
end
return clues

Algorithm 3: Clustering Strategy

Input : vectors, target_city, taboo_words, clusters
clues ← emptyList()
for cluster in clusters do
    clues_per_cluster ← emptyList()
    for example in cluster do
        clue ← performAnalogy(target_city, example)
        if allowed(clue, taboo_words) then
            clues_per_cluster ← add(clues_per_cluster, clue)
        end
        best_clue_from_cluster ← getHighestScoring(clues_per_cluster, target_city)
        clues ← add(clues, best_clue_from_cluster)
    end
end
return clues

5 Experiments and Results

In this section, the ADA architecture and its methods for the generation of word associations are evaluated through simulation in Subsection 5.1 and an empirical study in Subsection 5.2.

5.1 Automatic Evaluation

The AGA architecture and the Wiki* vectors presented by Dankers et al. (2017) were employed to evaluate the ADA through simulation. Firstly, the experimental setup is detailed and secondly, the results are presented.
Experimental Setup The ADAs that are evaluated employ four strategies: using taboo words, using taboo words and AE, applying clustering and analogy reasoning combined with IDF selection, and applying analogy reasoning with Cosine similarity selection. The second ADA employs the pre-trained Google News vectors. The first, third and fourth ADA employ the Wiki* vectors presented by Dankers et al. (2017) along with the Enumeration game strategy and Cosine similarity. The vectors used have 300 dimensions and were trained with a context-predicting DSM named Skip-Gram (Mikolov et al., 2013) and a context window size of 25. The vectors originally contained unigrams, but were retrained for the experiments conducted for this thesis to include bigrams and trigrams, considering that clues for the LTG are allowed to contain up to three words. The phrase detector used to implement this, is from the gensim.models.phrases\textsuperscript{15} library; only phrases that appeared at least five times were included in the vocabulary.

The experiments conducted serve to investigate the effect of adding the mechanism for dependent clues to the guesser agent and the effect of including bigrams and trigrams for the Wiki* vectors. The ADAs are evaluated according to their accuracy, which is the percentage of games in which the target city was guessed correctly. The cities used are from dataset1. The 100 cities used in the development of the AGA were excluded. The ADAs employing the Wiki* vectors are evaluated with an AGA employing the Google News vectors and vice versa to avoid bias in the results from both the AGA and the ADA employing the same vector space. Despite the application of POS-tagging to exclude clues that contain pronouns, manual inspection revealed that many pronouns were simply tagged as nouns, which may bias the results for the simulations conducted with the ADA and the AGA. For the empirical study, clues containing pronouns have been removed manually.

Results The results for the experiments conducted are displayed in Table 5. Although using only unigrams yields the highest accuracies, unigrams, bigrams and trigrams are included in the empirical study, to allow the ADA to give its users clues containing up to three words.

Table 5: An overview of the experiments conducted for the automatic evaluation, where Without and With concern the interactiveness enhancement for the AGA.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>AGA</th>
<th>ADA</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without</td>
<td>With</td>
<td>Without</td>
</tr>
<tr>
<td>Clustering, IDF</td>
<td>Google News</td>
<td>Wiki*, unigrams</td>
<td>21.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni- and bigrams</td>
<td>14.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni-, bi- and trigrams</td>
<td>11.97</td>
</tr>
<tr>
<td>Clustering, Cosine</td>
<td>Google News</td>
<td>Wiki*, unigrams</td>
<td>35.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni- and bigrams</td>
<td>20.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni-, bi- and trigrams</td>
<td>18.46</td>
</tr>
<tr>
<td>Taboo Words</td>
<td>Google News</td>
<td>Wiki*, unigrams</td>
<td>26.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni- and bigrams</td>
<td>23.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wiki*, uni-, bi- and trigrams</td>
<td>23.08</td>
</tr>
<tr>
<td>Taboo Words, AE</td>
<td>Wiki*, uni-,</td>
<td>Google News</td>
<td>26.50</td>
</tr>
<tr>
<td></td>
<td>bi- and trigrams</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{15}https://radimrehurek.com/gensim/models/phrases.html
5.2 Empirical Evaluation

To evaluate the describer agent empirically, an experiment has been designed that allows users to play the games directly with the describer agent. Firstly, the setup is explained and secondly, the results are presented.

**Experimental Setup** In the experiment, 36 games were made available for users, with 9 games for every strategy. These 36 games were distributed over 4 sessions, to avoid mood bias in the results. The first three sessions contained 3 games for the taboo words, the taboo words and AE and the clustering with idf strategies. The fourth session contained games for the clustering strategy, combined with the Cosine similarity. From dataset 1, cities from European countries were selected and ordered on popularity according to the ranking of NomadList\(^{16}\) and split into three equally sized groups, representing difficulty levels (DL) 1, 2 and 3. Every strategy received three cities from each category, to distribute the cities with differing guessing difficulty equally over the different strategies. An overview of the games per session is presented in Table 6. The maximum number of game iterations has been set to \(i = 12\). The number of available clues differs per city and per strategy.

The games were played by 9 participants, whose ages varied from 17 to 60. All subjects were born in the Netherlands, and all subjects were living in the Netherlands at the time of the experiment. Before playing the games, they filled out the opening questionnaire (Appendix A.1). The LTG rules, as explained in Section 3, were explained at the end of the opening questionnaire. After each game the participants filled out a few questions in the corresponding session questionnaire (Appendix A.2) and after completing the sessions, the closing questionnaire was filled out (Appendix A.3).

Table 6: An overview of the experimental setup, where each game has been given an identifier (ID) that corresponds to the results per game presented in Appendix B.1.

\(^{16}\)http://www.nomadlist.com
Results  In total, 27 sessions were completed by the participants, which resulted in a total of 243 games that were available for the preparation of the results and the analysis. The questionnaire for every game contained a question about whether the participant knew the city. To calculate how well-known the cities were, only participants who entered ‘Yes, but I do not know much about it.’ or ‘Yes, I know the city very well.’ were counted as knowing the city. These participants were further asked to enter their rating of the description and their results were used to calculate the average accuracy. On average, participants indicated that they knew 80% of the cities that were labelled ‘easy’ (DL 1), 55.95% of the cities that were labelled ‘medium’ (DL 2) and 56.19% of the cities that were labelled ‘difficult’ (DL 3).

The average accuracy and average rating are displayed in Figure 2, for the three DLs and the different variants of strategies employed by the ADAs; these results are summarised in Table 7. In Appendix B.1 a more detailed overview of the results per game is presented.

Table 8 presents the results for questions from the opening and closing questionnaire, regarding the users’ topographical and geographical knowledge and their general evaluation of the ADAs. The results per participant are displayed in Appendix B, Table 10. Pearson’ correlation coefficient, $r$, can be used to investigate the relations between these different variables: The subjects’ ratings of their European topographical knowledge has a slightly positive correlation to the overall accuracy ($r = 0.131$) and a negative correlation to the accuracy for known cities ($r = -0.276$). The subjects’ ratings for their ability to recognise European cities from short descriptions has a positive correlation to the overall accuracy ($r = 0.602$) and a negative correlation to the accuracy for known cities ($r = -0.452$).

Figure 2: The average accuracy for known cities and average rating from the empirical study, represented per DL and per strategy.

Table 7: An overview of the overall accuracy for the games in which the target city was known and the average rating per strategy variant.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Accuracy (%)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>TabooWords</td>
<td>59.52</td>
<td>6.89</td>
</tr>
<tr>
<td>TabooWords, AE</td>
<td>47.72</td>
<td>6.40</td>
</tr>
<tr>
<td>Clustering, IDF</td>
<td>42.85</td>
<td>4.98</td>
</tr>
<tr>
<td>Clustering, Cosine</td>
<td>54.84</td>
<td>6.54</td>
</tr>
</tbody>
</table>

16
6 Discussion

In this work, an ADA architecture has been presented that aims at generating associations with cities, to which humans can relate. Several variants of this architecture have been evaluated through simulation and an empirical study, in which 9 subjects participated. The subjects were able to guess the city correctly for 37.86% of the games and for 50.00% of the games in which the guesser knew the target city (Research Question 1). The language used by the ADA has been rated with a 5.73, which illustrates that the clues do not entirely resemble language humans would use. The target cities with different guessing difficulties were equally distributed over the strategies and sessions. The results suggest that level 1 cities were indeed the easiest cities to guess and were assigned the highest rating on average, as can be seen in Figure 2. Through the questionnaires filled out by the participants their topographical and geographical knowledge could be related to their performance in the game. There was a positive correlation between their knowledge in these area’s and their performance on all games. However, there was a negative correlation with the performance of games for which the target city was known by the guesser. This negative correlation may be explained by the number of candidate cities, which depends on a subject’s topographical and geographical knowledge. More knowledge in these areas affects the performance for all known cities negatively a priori, considering that the subject has more cities to choose from. The clues that the participants found most helpful, were clues about food. This illustrates that a clue’s quality does not correspond to the similarity between a clue and the target city in the vector space, as can be seen in Figure 1.

Two methods have been implemented to extract word associations from the vector space (Research Question 2). The first method uses the nearest neighbours of taboo words, which has been combined with the AE algorithm. The second method performs analogy reasoning with clusters of data from games played by humans, selecting one clue per cluster for a target city. This method has been combined with two measures for the selection of the optimal clue per cluster. The results from both the simulation and the empirical study suggest that using the taboo words results in a better performance on average. Applying the AE algorithm did not improve the performance. However, relying on the availability of taboo words implies that the ADA does not generate entirely new associations for the target city itself, considering that the lists of taboo words have been put together by humans. The clustering approach moves away from the specific rules of the LTG and is more general, since it can generate associations for a target city for which no associations prepared by humans are available. The results for both evaluations suggest that selecting clues per clusters according to the Cosine similarity to the target city, yields higher accuracies than applying the IDF. A disadvantage of the clustering...
strategy for the preparation of independent clues, is that for every city a fixed number of clues is prepared in fixed categories, which are categories such as food, climate or history. Although this might result in a recognisable description for many cities, there are exceptions that are only well-known because of one characteristic. For example, Dresden is predominantly well-known because of the bombing of Dresden and Cannes is well-known because of the yearly film festival. Although these topics might be covered through analogies for the clusters of ‘History’ and ‘Media & Culture’, the description as a whole would differ from how humans would describe these cities and the added information from the other clusters might put the guesser off. For cities like these, the clustering approach might be less effective.

A rule-based approach has been presented that serves to create clues that depend upon the guess given by the user (Research Question 3), allowing the describer agent to respond with a cardinal direction, ‘close’, ‘different country’ or ‘different continent’. Through these dependent clues, the ADA can display collaborative behaviour towards the guesser, by guiding him into the right direction, or away from the wrong direction. The collaborative behaviour has been rated by the participants with a 5.90 on average. One of the participants entered ‘different country’ as one of the types of clues that he found most helpful. This illustrates that the ADA succeeded in giving useful dependent clues. Nevertheless, this approach could be improved by a more advanced mechanism for deciding when to respond with a dependent guess. One participant commented on a game that the ADA waited five iterations before mentioning that the participant was guessing cities situated on the wrong continent. Also, for participants with little topographical knowledge responding with cardinal directions may not bring the guesser closer to the right answer.

7 Conclusion and Future Work

In this thesis, multiple methods for extracting word associations from a vector space have been presented and applied to the creation of an ADA for the LTG. Those methods employ a nearest neighbours approach by utilising the taboo words per game or apply clustering and analogical reasoning to data of games played before. These ADAs have been evaluated through simulation and an empirical study and could elicit a correct guess for 37.86% of the games played, and for 50.00% of the games in which the participants knew the target city. Using the taboo words is a method that is specific to the LTG, but the clustering approach can be generalised and applied to different types of word-guessing games. A rule-based approach for modelling interactiveness has been proposed that allows the describer agent to present clues that depend upon the responses of the guesser, by expressing a relative geographical relation. The collaborative behaviour of the guesser has been valued with a 5.90 on a scale from 1 to 10.

Firstly, future research could focus on improving the dependent clues, by modelling their helpfulness for the user and by adding different types of dependent clues – e.g. ‘older city’ and ‘larger’. The number of dependent clues could be made dependent on whether they help the user or not, through a dynamic limit instead of a fixed probability.

Secondly, the methods presented for the generation of word associations can be improved upon. One aspect where improvement is needed is the order in which the clues are presented to the user. The order is currently based upon the clues’ Cosine similarity to the target city in the vector space, but this order does not necessarily present the clue that’s the most useful for the user first. For example, although ‘boat tours’ is generally regarded as a strong association for Venice, this clue was ranked eighth in the description generated by the strategy that uses taboo words and the AE algorithm.
Thirdly, future work could include modelling individual word spaces that better capture an individual’s knowledge. Human associations are sensitive to many factors, such as one’s economic and social status, one’s country or place of residence, political and religious conviction and cultural background (Sahlgren, 2006). Although the vector spaces used in this work may represent perfectly associations that follow from the corpora used, they may not be representative of the personal word spaces the test subjects have. To be able to judge more accurately whether the describer agents employs a good associative strategy, a word space more representative of the knowledge of test subjects could be employed.

References


Dankers, V. (2017). Modelling the generation and retrieval of word associations with word embeddings (Bachelor’s Thesis). University of Amsterdam.


Marti, S. & Emnett, K. (n.d.). Daboo: An interactive system to make a user guess a word as fast as possible (without using taboo words).


A Questionnaire

A.1 Opening Questionnaire

Travel History

1. How many countries did you visit within Europe?
   (a) 1-5
   (b) 5-10
   (c) 10-20
   (d) 20+

2. How many countries did you visit worldwide (Europe excluded)?
   (a) I have not traveled outside Europe.
   (b) 1-5
   (c) 5-10
   (d) 10-20
   (e) 20-50
   (f) 50+

Geographical Knowledge

3. On a scale from 1 to 10, how would you rate your topographical knowledge for Europe?
4. On a scale from 1 to 10, could you recognise Europe’s 25 most well-known cities from descriptions of just 20 words per city?
5. On a scale from 1 to 10, how would you rate your topographical knowledge worldwide (Europe included)?
6. On a scale from 1 to 10, could you recognise the world’s 100 most well-known cities from descriptions of just 20 words per city (European cities included)?

A.2 Questions per Game

1. Did you know the target city?
   (a) No / I have only heard of its name.
   (b) Yes, but I do not know much about it.
   (c) Yes, I know the city very well.

2. If you knew the target city, how well was the description of the describer agent in hindsight on a scale from 1 to 10?
3. Do you have any comments about the game?
A.3 Closing Questionnaire

Location Taboo Game

1. On a scale from 1 to 10, how much did you enjoy the game?

2. On a scale from 1 to 10, how well did the language used resemble natural language humans would use in word-guessing games (if restricted according to the rules specified for the Location Taboo Game)?

3. On a scale from 1 to 10, how would you rate the behaviour of the describer as your teammate?

4. Did you cheat during the game?

5. What types of clues do you find most useful? Select three options at most.
   (a) Clues about food (for example ‘pasta’ for Italian cities)
   (b) Clues about history (for example ‘bombing’ Dresden or Rotterdam)
   (c) Clues about weather or environment (for example ‘very cold’ for Reykjavik)
   (d) Clues about things you’d visit as a tourist (for example ‘bell tower’ for Pisa)
   (e) Clues about culture (for example ‘dance’ for Havana)
   (f) Other

Demographics

6. What is your age?

7. What is the highest degree or level of education you have completed?

8. In which country do you live?

9. In which country were you born?
**B Empirical Evaluation**

**B.1 Results per Game**

Table 9: An overview of the results per game. The familiarity represents the percentage of people who entered ‘Yes, but I do not know much about it.’ or ‘Yes, I know the city very well.’ as an answer to the question ‘Did you know the target city?’. Only the results of those participants who knew the city were used for the average accuracies and average ratings of the descriptions.

<table>
<thead>
<tr>
<th>City</th>
<th>Average Indicated Familiarity with Target City (%)</th>
<th>Accuracy (%)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athens</td>
<td>100.00</td>
<td>100.00</td>
<td>6.86</td>
</tr>
<tr>
<td>Bratislava</td>
<td>57.14</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Sarajevo</td>
<td>57.14</td>
<td>0.00</td>
<td>5.80</td>
</tr>
<tr>
<td>Prague</td>
<td>85.71</td>
<td>14.29</td>
<td>5.33</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>71.43</td>
<td>60.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Dresden</td>
<td>71.43</td>
<td>0.00</td>
<td>3.67</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>100.00</td>
<td>100.00</td>
<td>5.57</td>
</tr>
<tr>
<td>Naples</td>
<td>100.00</td>
<td>0.00</td>
<td>5.14</td>
</tr>
<tr>
<td>Strasbourg</td>
<td>85.71</td>
<td>16.67</td>
<td>7.00</td>
</tr>
<tr>
<td>Paris</td>
<td>100.00</td>
<td>100.00</td>
<td>8.71</td>
</tr>
<tr>
<td>Bordeaux</td>
<td>57.14</td>
<td>75.00</td>
<td>8.75</td>
</tr>
<tr>
<td>Ankara</td>
<td>28.57</td>
<td>50.00</td>
<td>5.50</td>
</tr>
<tr>
<td>Barcelona</td>
<td>85.71</td>
<td>100.00</td>
<td>8.17</td>
</tr>
<tr>
<td>Florence</td>
<td>42.86</td>
<td>100.00</td>
<td>8.33</td>
</tr>
<tr>
<td>Birmingham</td>
<td>42.86</td>
<td>33.33</td>
<td>6.00</td>
</tr>
<tr>
<td>Lisbon</td>
<td>28.57</td>
<td>50.00</td>
<td>6.50</td>
</tr>
<tr>
<td>Nottingham</td>
<td>14.29</td>
<td>100.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Turin</td>
<td>71.43</td>
<td>50.00</td>
<td>7.20</td>
</tr>
<tr>
<td>Budapest</td>
<td>42.86</td>
<td>33.33</td>
<td>6.00</td>
</tr>
<tr>
<td>Manchester</td>
<td>42.86</td>
<td>33.33</td>
<td>6.50</td>
</tr>
<tr>
<td>Bern</td>
<td>57.14</td>
<td>50.00</td>
<td>6.25</td>
</tr>
<tr>
<td>Berlin</td>
<td>85.71</td>
<td>50.00</td>
<td>5.83</td>
</tr>
<tr>
<td>Marseille</td>
<td>42.86</td>
<td>66.67</td>
<td>5.33</td>
</tr>
<tr>
<td>Kiev</td>
<td>28.57</td>
<td>0.00</td>
<td>2.50</td>
</tr>
<tr>
<td>Madrid</td>
<td>71.43</td>
<td>100.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Sofia</td>
<td>42.86</td>
<td>0.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Verona</td>
<td>71.43</td>
<td>20.00</td>
<td>5.80</td>
</tr>
<tr>
<td>Brussels</td>
<td>80.00</td>
<td>100.00</td>
<td>6.50</td>
</tr>
<tr>
<td>Lyon</td>
<td>60.00</td>
<td>33.33</td>
<td>5.67</td>
</tr>
<tr>
<td>Oxford</td>
<td>80.00</td>
<td>33.33</td>
<td>6.75</td>
</tr>
<tr>
<td>Vienna</td>
<td>100.00</td>
<td>60.00</td>
<td>6.20</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>100.00</td>
<td>20.00</td>
<td>7.00</td>
</tr>
<tr>
<td>Palermo</td>
<td>40.00</td>
<td>60.00</td>
<td>8.50</td>
</tr>
<tr>
<td>Venice</td>
<td>80.00</td>
<td>50.00</td>
<td>6.25</td>
</tr>
<tr>
<td>Innsbruck</td>
<td>40.00</td>
<td>0.00</td>
<td>4.50</td>
</tr>
<tr>
<td>Aberdeen</td>
<td>40.00</td>
<td>50.00</td>
<td>6.00</td>
</tr>
</tbody>
</table>
### B.2 Results per Participant

Table 10: Game results per participant. The accuracies and familiarity are averages for the games played for that participant, which differs per participant. A.1.3 and A.1.4 represent the answers to the questions 3 and 4 from the opening questionnaire (Appendix A.1).

<table>
<thead>
<tr>
<th>Participant</th>
<th>Overall Accuracy (%)</th>
<th>Accuracy for Known Cities (%)</th>
<th>Average Indicated Familiarity with Target Cities (%)</th>
<th>A.1.3</th>
<th>A.1.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>F3</td>
<td><strong>52.78</strong></td>
<td>55.88</td>
<td><strong>94.44</strong></td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>F6</td>
<td>47.22</td>
<td>47.06</td>
<td><strong>94.44</strong></td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>F11</td>
<td>44.44</td>
<td>66.67</td>
<td>66.67</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>F10</td>
<td>38.89</td>
<td>50.00</td>
<td>44.44</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>A1</td>
<td>33.33</td>
<td>42.86</td>
<td>77.78</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>F2</td>
<td>33.33</td>
<td><strong>100.00</strong></td>
<td>22.22</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>F1</td>
<td>33.33</td>
<td>50.00</td>
<td>61.11</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>F5</td>
<td>29.63</td>
<td>40.00</td>
<td>74.07</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>F4</td>
<td>25.00</td>
<td>41.18</td>
<td>44.44</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>