ASCILITE
25-28 November 2018
Open Oceans:
Learning without borders

ASCILITE 2018

Australasian Society for Computers
in Learning in Tertiary Education
Deakin University, Geelong, Australia

Conference Proceedings

25–28 November 2018

Open Oceans:
Learning Without Borders

35th International Conference
of Innovation, Practice and Research
in the use of Educational Technologies
in Tertiary Education

DEAKIN UNIVERSITY

ASCILITE
SHAPING THE FUTURE OF TERTIARY EDUCATION

ASCILITE 2018

Australasian Society for Computers in Learning in Tertiary Education

Deakin University, Geelong, Australia
25-28 NOVEMBER 2018

Conference Proceedings

OPEN OCEANS: LEARNING WITHOUT BORDERS

35TH INTERNATIONAL CONFERENCE OF INNOVATION, PRACTICE AND RESEARCH IN THE USE OF EDUCATIONAL TECHNOLOGIES IN TERTIARY EDUCATION

Editors:
Malcolm Campbell, Julie Willems, Chie Adachi, Damian Blake, Iain Doherty, Siva Krishnan, Susie Macfarlane, Leanne Ngo, Marcus O'Donnell, Stuart Palmer, Lynn Riddell, Ian Story, Harsh Suri & Joanna Tai

Parts of speech in Bloom's Taxonomy Classification

Brian R. von Konsky
Curtin University,
Australia

Longwei Zheng
East China Normal University,
China

Eric Parkin
Curtin University
Australia

Simon Huband
Curtin University,
Australia

David C. Gibson
Curtin University,
Australia

This paper analyses parts of speech in a training corpus with 13,189 learning outcomes in which Bloom's Taxonomy levels were previously classified by human experts for 3,496 subjects offered at an Australian university. This paper explores the automatic identification of verbs and other parts of speech impacting the semantic meaning and Bloom's classification of learning outcome statements. The frequency with which words in learning outcomes appear as different parts of speech and at different Bloom's levels is described as a preliminary step of a larger project that aims to automatically classify Bloom's levels using a combination of table lookup and machine learning approaches. It is indicated that automated parts of speech classification can assist human learning and teaching designers to write clearer learning outcome statements. This is in addition to playing a role in automated Bloom's Taxonomy classification, and identifying cases requiring review in conjunction with normal institutional curriculum management processes.

Keywords: Bloom's Taxonomy, Learning Outcomes, Machine Learning, Parts of Speech.

Introduction

Bloom's Taxonomy is widely used as a means of describing the level of cognition expected in student learning activities and assessments (Bloom, Kratwohl, & Masia, 1956). Table 1 shows the 6 hierarchical levels of the revised Bloom's taxonomy (Anderson & Krathwohl, 2001), the meaning associated with each level, and a partial list of indicative verbs that can be used to classify associated learning outcomes.

Table 1: Bloom's Taxonomy

Bloom's level	Semantic meaning	Indicative verbs
1. Remembering	Simple recall often associated with memorisation	list, name, state, define
2. Comprehending	Basic understanding sufficient to explain ideas to others	identify, explain, describe
3. Applying	Use of information or knowledge in new ways	apply, use, solve, compute
4. Analysing	Establishing connections or relationships	analyse, compare, classify
5. Evaluating	Form a judgement or critique	evaluate, appraise
6. Creating	Synthesising something new	create, design, plan, compose

Stanny (2016) conducted a meta-analysis that identified Bloom's Taxonomy verbs from 30 sources. Her analysis considered verbs and the frequency with which they were included in verb tables available online. She included a verb in a conservative aggregation if a verb occurred at a given Bloom's level in 10 or more of the 30 tables included in the study. This resulted in a table with 104 unique verbs, out of 128 verbs in total. That is, some verbs can be indicative of more than one Bloom's category. For example, the verb *identify* can be indicative of the *comprehending*, *applying*, or *analysing* category depending on the context in which it is used.

Machine learning and probabilistic parsers can automatically identify parts of speech in an arbitrary sentence (Klein & Manning, 2002, 2003; Müller & Guido, 2016). In such an approach, parts of speech are determined using a model based on a training corpus in which a human expert has previously tagged parts of speech. Once identified, verbs can be used in lookup tables to identify indicative Bloom's categories. In those instances in which a verb can indicate different Bloom's categories, Omar et al. (2012) assigned a weight determined by subject matter experts in a rule based approach to determine the likely classification. Similarly, Yahya, Sman, Taleb, and Alattab (2013) have used Machine Learning to classify cognition levels based on training data in which human experts had previously assigned Bloom's level to classroom questions.

Objectives and Methodology

This study used automated parts of speech tagging and verb table lookup to: 1) explore how the approach can assist human experts to write learning outcome statements that clearly articulate what is expected of students and that are free of grammatical ambiguity; 2) examine how parts of speech other than verbs impact outcome semantics; and 3) serve as a base-lined for subsequent Machine Learning based Bloom's classification.

A total of 13,189 learning outcomes from all undergraduate and postgraduate subjects were downloaded from the curriculum database of an Australian university. There were 8115 learning outcomes from undergraduate subjects and 5074 were from postgraduate subjects. The University's central teaching organisation had previously participated in the Bloom's classification of each learning outcome as part of the institution's curriculum management process. The distribution of learning outcomes by Bloom's level is shown in Table 2. There were relatively few examples of *Remembering*. In undergraduate subjects, there were more examples of *Applying* than any other Bloom's category (24.6%, *N*=1996). For postgraduate subjects, there were more examples of Creating than other categories (31.3%, N=1598), followed closely by Evaluating (30.2%, N=1533).

Table 2: The distribution of Bloom's levels for undergraduate and postgraduate subjects.

Bloom's Classification	Undergraduate (%)	Undergraduate (N)	Postgraduate (%)	Postgraduate (N)
Remembering	1.36	110	0.67	34
Comprehending	11.61	942	5.22	265
Applying	24.60	1996	15.16	769
Analysing	19.57	1588	17.42	884
Evaluating	21.84	1772	30.21	1533
Creating	21.04	1707	31.32	1589

Parts of speech were automatically identified for each outcome statement using a public domain parser from Stanford University in a Python program using the Natural Language ToolKit (NLTK) package. A pre-processing step appended each learning outcome to: "On successful completion of this unit students can". This was done to form grammatically correct sentences before processing. From this, a feature set consisting of all 7929 unique words in the learning outcome corpus was constructed. This recorded the number of times that each word was categorised as a given part of speech or appeared in an outcome statement at a given Bloom's level.

Verbs were automatically identified to determine the indicated Bloom's level using a lookup table based on the full meta-analysis by Stanny (2016). In those instances where a verb appeared in the lookup table in multiple Bloom's categories, the level with the highest corpus frequency was selected. Outcome statements were classified using the highest cognition level indicated by identified verbs. The Bloom's level of each outcome statement in the corpus was classified using this approach and compared to the classification made in conjunction with prior institutional curriculum management processes and recorded in the University database.

Results

Figure 1 shows the accuracy of predicting the Bloom's level using the verb table lookup approach.

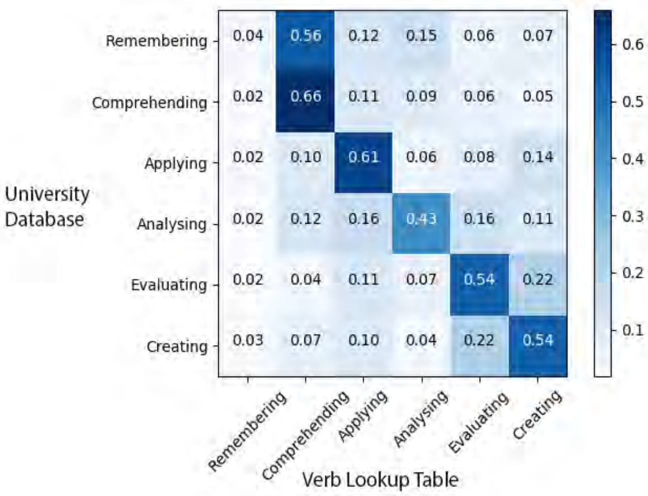

Figure 1: The accuracy of classifying the Bloom's level using the verb table lookup.

The figure shows the fraction of times that the Bloom's level as classified by humans and recorded in the University database was consistent with classification using automated verb table lookup. The goal is for high accuracy shown by darker cells along the diagonal axis of the matrix. Of the 13,189 outcome statements, the verb table lookup approach identified the same Bloom's level that was recorded in the University database for 55% (*N*=7081) of the outcome statements.

Results suggest that the verb table lookup approach had difficulty in differentiating between the *remembering* and *comprehending* Bloom's levels, with 55.6% (*N*=80) of the *remembering* outcome statements being identified as *comprehending*. There were 22.4 % (*N*=742) of the *evaluating* outcome statements that were identified as *creating*, and 22.2% (*N*=732) of *creating* cases were identified as *evaluating*. There were 6.9% (*N*=10) instances of *remembering* outcome statements that were categorised as examples of *creating*. Similarly, 2.7% (*N*=92) instances of *creating* were categorised at the *remembering* Bloom's level.

Potential sources of difference between the actual and predicted Bloom's levels include: 1) failing to consider the semantic impact associated with parts of speech other than verbs; 2) parser errors associated with identifying pats of speech; 3) verbs missing in the lookup table; and 4) tacit knowledge about assessments and errors impacting the original classification.

As will be shown, classification discrepancies can be used to flag outcome statements for review as part of an institution's normal curriculum management process.

For example, an outcome statement that was recorded in the University database as *remembering*, but which was classified by as *creating* was: "develop understanding of the concepts of electronic devices and circuits." The verb *develop* is in the verb table as an example of *creating* because it is usually used in the context of synthesising something new. The word *understanding* is a noun and so it was not considered in the verb table classification. Moreover, this statement should be rewritten, as it does not say what the student must do to demonstrate that understanding has been developed.

Similarly, the outcome statement: "prepare management accounting data" is listed in the University database as being an example of the *applying*. Tacit knowledge about how this outcome is assessed may reasonably lead one to conclude that this is an example of applying a basic accounting management skill. Automatic table lookup, however, classifies this as an example of *creating* because *prepare* is at that level in the lookup table and can reasonably suggests that something new is being synthesised. Based on this tacit knowledge, this outcome statement might be left unaltered after review.

Note that identifying parts of speech is necessary when using automated table lookup because some verbs can also be nouns. That is, the presence in the lookup table of a word from the outcome statement by itself is insufficient for classification. Table 3 shows a portion of the feature set for the five most commonly occurring words in the corpus that were tagged as both verbs and nouns, but not other parts of speech.

Table 3: The five most frequently occurring words classified as both verbs and nouns

Feature	Verb	Noun	Adverb	Adjective	Other	Sum	%LO
research	111	1280	0	0	0	1391	10.5
design	407	650	0	0	0	1057	8.0
practice	29	796	0	0	0	825	6.3
use	371	201	0	0	0	572	4.3
work	198	280	0	0	0	478	3.6

In some cases, inspection of the tree produced by the parser demonstrated an accurate grammatical interpretation of the outcome statement, but with a meaning other than the one intended. For example, consider the outcome statement: "analyse design decisions and report findings". The intention had been that students would report on the findings of an analysis. The tree produced by the parser for this outcome statement is shown in Figure 2. The parser identified a verb phrase (VP) consisting of the verb (VB) *analyse* and a noun phrase (NP) with two parts combined by a coordinating conjunction (CC). The first part of the noun phrase consisted of the singular noun (NN) *design* and the plural noun (NNS) *decision.* The second part consisted of the singular noun (NN) *report* and the plural noun (NNS) *findings*. The parser tagged *report* as a noun rather than verb and interpreted the outcome statement to mean that both design decisions and report findings were to be analysed.

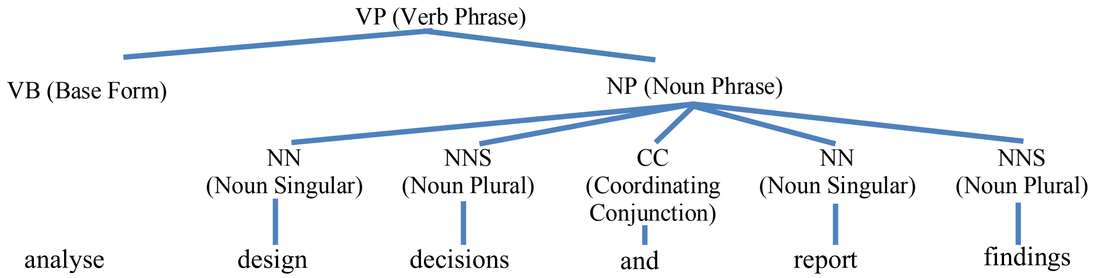

Figure 2. Parse tree for "analyse design decisions and report findings"

Confronted with this interpretation, a learning designer might choose to rewrite these as two separate outcome statements, or rewrite the statement to avoid the ambiguity. For example, changing the outcome statement to "analyse design decisions and then report findings" corrects the ambiguity and the statement parses as expected.

Not surprisingly, *apply*, *evaluate*, *analyse*, *explain*, and *demonstrate* were the 5 most commonly occurring verbs in the feature set, with frequency counts show in Table 4. The table shows the number of times that these features appeared in outcome statements that were tagged at a given Bloom's level, with the maximum value for each feature being shown with a blue background. Outcome statements often contain more than one verb, so the frequency does not necessarily indicate that a feature is representative of that Bloom's level. However, the maximum frequency for each feature is generally consistent with expectations. The frequency with which features were tagged in outcome statements at given part of speech is also shown. The maximum value is shown with an orange background.

Table 4. The 5 most frequently occurring verbs.

Feature	Occurrences as a given Bloom's Level						Occurrences as a given part of speech					Sum
	Remembering	Comprehending	Applying	Analysing	Evaluating	Creating	Verb	Noun	Adjective	Adverb	Other	
apply	7	44	1094	295	291	222	1953	0	0	0	0	1953
evaluate	2	12	58	81	1021	503	1662	0	15	0	0	1677
analyse	1	21	86	966	271	208	1542	0	10	1	0	1553
explain	12	366	167	327	80	42	994	0	0	0	0	994
demonstrate	3	43	461	95	121	112	831	0	4	0	0	835

The verbs in Table 4 were tagged as an invalid part of speech less than 1% of the time. That is, the verb *evaluate* was misclassified as a adjective 0.8% (N=15) of the 1677 times it occurred in as a feature in the corpus. The verb *analyse* was misclassified as either an adjective or adverb 0.7% of the 1553 times it occurred in the corpus (Adjective, N=10; Adverb, N=1). The verb *demonstrate* was misclassified as an adjective 0.5% (N=4) of the 835 times it occurred in the corpus.

Inspection showed that misclassification sometimes occurred in complex sentences containing adjectives where those adjectives are words that also have a verb form in other contexts. For example, misclassification occurred for "critically evaluate food processing unit operations and related equipment". In this example, *related* should have been classified as an adjective, *critically* as an adverb, and *evaluate* as a verb. In other contexts, however, *related* could form the past participle of the verb *relate* and the parser failed to correctly tag the outcome. Less complex sentences with the adjective *related* were seen to parse correctly.

In this example, the adverb *critically* could be removed with no impact on statement semantics or the resulting Bloom's classification. This is because the verb *evaluate* is already indicative of the *evaluating* Bloom's category, which expects the student to form an opinion or make a judgement. That is, "evaluate food processing unit operations and related equipment" would be classified at the same Bloom's level as the original statement. The meaning and Bloom's classification is not changed by removing the adverb *critically* in this case.
This is different than situations in which the adverb *critically* is used in conjunction with the verb *analyse*. Consider the following outcome statement: "critically analyse the characteristics of different industry sectors and explain a firm's competitive strategy". The University database identifies this as an example of the

evaluating Bloom's category. That is, to *critically analyse* something is different than just *analysing* it because the student is being asked to critique the analysis and form a judgement and do more than just establish connections or relationships. In this instance, the adverb *critically* has changed the semantic meaning of the outcome statement and hence the resulting Bloom's classification. This suggests that the verb *evaluate* may be a better choice for the outcome statement over the semantically equivalent *critically analyse*.

Similarly, the adjective *significant* in "analyse data and communicate significant findings" calls for the student to make a judgement about findings that have been analysed. Although the verb analyse might suggest that this outcome statement is an example of the *analysing* Bloom's level, the adjective *significant* has changed the classification to make this an example of *evaluating*.

The adverb appearing most frequently in outcome statements was *critically* (N=762). The frequency with which the *adverb* critically appeared in outcome statements at each Bloom's levels were: remembering, 0% (N=0); understanding, 2.2% (N=17); applying, 2.4% (N=18); analysing, 20.8% (N=159); evaluating, 46.6% (N=355); and creating, 28.0% (N=213). Excluding simple adverbs like *how*, *when*, *as*, and *well*, other adverbs in decreasing order of frequency and with 20 or more occurrences were *effectively* (N=300), *independently* (N=70), *appropriately* (N=65), *professionally* (N=64), *clearly* (N=44), *collaboratively* (N=43), *culturally* (N=43), *internationally* (N=30), *orally* (N=29), *safely* (N=25), *commonly* (N=22), and *accurately* (N=20). Occurring only once, the adverb *innovatively* was used in the outcome statement: "innovatively apply knowledge and skills, techniques and methods to the process of studio practice…" The university database identified this as an example of the *applying* Bloom's level. However, to innovatively apply knowledge suggests that the student is being asked to do something new that has not been done before. As such, this outcome statement is arguably an example of the *creating* Bloom's level because of this adverb.

Conclusion

Automated parts of speech identification and verb table lookup provides a means to automatically classify Bloom's outcome statements. This can be used to identify statements for review in the context of institutional curriculum management processes, or to assist in writing clear outcome statements that are free from ambiguity. The verb table approach will serve as the baseline for a subsequent Machine Learning approach to Bloom's classification that is currently under investigation, which will include words other than verbs in the training data. As shown in this paper, although not widely discussed in the literature, parts of speech other than verbs can impact the meaning of a learning outcome statement and the resulting Bloom's classification. As such, it is anticipated that a Machine Learning approach will improve the accuracy of Bloom's classification.

Acknowledgements

This research was supported in part by Curtin University and the Higher Education Text Analytics (HETA) project funded by a Learning and Teaching Grant from the Australian Technology Network (ATN).

References

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Bloom, B. S., Kratwohl, D. R., & Masia, B. B. (1956). *Taxonomy of educational objectives: The classification of educational goals*. London: McKay.

Klein, D., & Manning, C. D. (2002). *Fast exact inference with a factored model for natural language parsing*. Proceedings of the 15th International Conference on Neural Information Processing Systems.

Klein, D., & Manning, C. D. (2003). *Fast exact inference with a factored model for Natural Language Parsing.* Proceedings of the 41st Meeting of the Association for Computational Linguistics.

Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning in Python: A guide for Data Scientists*: O'Reilly Media, Inc.

Omar, N., Haris, S. S., Hassan, R., Arshad, H., Rahmat, M., Zainal, N. F. A., & Zulkifli, R. (2012). Automated analysis of exam questions according to Bloom's taxonomy. *Social and Behavioral Sciences, 59*, 297-303.

Stanny, C. J. (2016). Reevaluating Bloom's Taxonomy: What measurable verbs can and cannot say about student learning. *Education Science, 6*(37), 12.

Yahya, A. A., Sman, A., Taleb, A., & Alattab, A. A. (2013). Analyzing the cognitive level of classroom questions using machine learning techniques. *Procedia - Social and Behanvioral Sciences, 97*, 587-596.

Please cite as: von Konsky, B.R., Zheng, L., Parkin, E., Huband, S. & Gibson, D. (2018). Parts of speech in Bloom's Taxonomy Classification. In M. Campbell, J. Willems, C. Adachi, D. Blake, I. Doherty, S. Krishnan, S. Macfarlane, L. Ngo, M. O'Donnell, S. Palmer, L. Riddell, I. Story, H. Suri & J. Tai (Eds.), Open Oceans: Learning without borders. Proceedings ASCILITE 2018 Geelong (pp. 527-532).