Teaching NLP/CL through Games: the Case of Parsing

Hans van Halteren
Department of Language and Speech
University of Nijmegen
P.O.Box 9103, NL-6500 HD, Nijmegen, The Netherlands
hvh@let.kun.nl

Abstract

This paper advocates the use of games in teaching NLP/CL in cases where computational experiments are impossible because the students lack the necessary skills. To show the viability of this approach, three games are described which together teach students about the parsing process. The paper also shows how the specific game formats and rules can be tuned to the teaching goals and situations, thus opening the way to the creation of further teaching games.

1   Introduction

Experience is the best teacher, as proverbial wisdom tells us. This should bode well for teaching NLP/CL, where the implementation is normally as important as the theory, and students can get hands-on experience with most of the field. They can use existing NLP/CL systems, seeing what inputs can be dealt with, what outputs are produced, and what the effects of different parameters settings are. They can even (re)implement parts of systems, or whole systems of their own. So, at a first glance NLP/CL appears to be the ideal field for teaching by means of experiments. However, there are also circumstances which make personal experimentation impossible. One obvious limitation is shortage of time. Experiments, especially ones in which a system is thoroughly examined, tend to take up a prohibitive amount of time for most course schedules. This can be worked around by focusing the experiments on the more important and/or more widely instructive aspects. Harder to work around is the situation where experiments are impossible because the students

lack the necessary knowledge. Students cannot program a system if they have insufficient programming skills; they cannot alter computational grammars if they have insufficient linguistic skills. This problem typically occurs in introductory courses, e.g. for general linguistics students, where the necessary knowledge is acquired only later (or never at all). If we want to keep the added value of personal experience, we will have to cast the intended experiencing into a different form.

In this paper I propose the use of games, another time-honored learning method, even for linguistics (cf. Grammatical Game in Rhyme (A Lady, 1802), teaching parts of speech, or, more recently, WFF 'N PROOF games like Queries 'n Theories (Allen et al., 1970), teaching formal grammars). I will not go into computer games or simulations (e.g. VISPER; Nouza et al., 1997), assuming these to be sufficiently known, but will focus on three major types of 'unplugged' games instead: card games, board games and roleplaying games. In the following sections, I give an example of each type. Together, these examples form an introduction to parsing and parsers, showing what kind of grammatical units are used (card game), how sentences can be broken down into these units by following a recursive transition network (RTN; board game) and how a parser can decide which route to take within the RTN (roleplaying game).

In all three games I use the descriptive model underlying the TOSCA/ICE parser (cf. Oostdijk, 2000), which in turn took its inspiration from the widely used descriptive grammars of Quirk et al. (1972, 1985). This is a constituent structure model where each constituent is labelled with a syntactic category (signifying what type of constituent it is) and a syntactic function (signifying what the constituent's role is in the

immediately dominating constituent).[1] Furthermore, all utterances and analyses in the games are taken from actual syntactically analysed corpus material, to be exact from a single 20,000 word sample taken from a crime novel (Allingham, 1965). This choice of text material and descriptive model is not just made out of convenience. I feel that it is important that the students work with 'real' examples, and not with especially constructed sentences and/or toy grammars.

2 A Card Game on Syntactic Building Blocks

The introduction to parsing starts with a card game about syntactic constituents and their interrelations. After all, if we want the students to understand what a parser does, they will first have to learn about the building blocks that are used in syntactic analysis. Even if they have already taken a syntax course, it will be necessary to familiarize them with the specific grammatical units used in our own 'parser' (i.e. those from the TOSCA/ICE model). As the main goal is familiarization with the terminology, we do not want to spend too much time on this. Also, this may be the students' first encounter with syntactic analysis, so they should be able to focus on the sentences and not be distracted by game rules. These two demands lead us to create a short (half hour) card game with simple rummy-like rules, Ling Rummy.

The Ling Rummy deck consists of 54 cards,[2] each of which (see Figure 1) depicts a syntactic function (e.g. CO = object complement), a terminal syntactic category (e.g. ART = article), a non-terminal syntactic category (e.g. NP = noun phrase), and an utterance. The goal of the game is to form combinations of three cards, a constituent in the utterance shown on the first card having a category shown on the second and the function shown on the third card (e.g. in Figure 1 "absolutely quiet" is an adjective phrase (AJP) functioning as an object
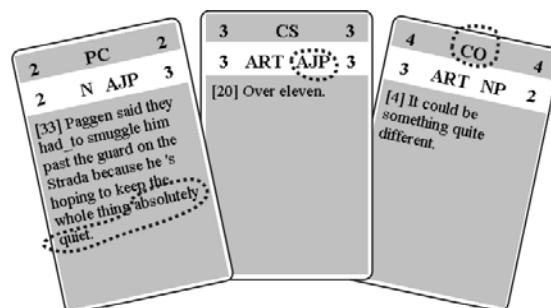

Figure 1: Three game cards from Ling Rummy with a scoring combination

complement (CO)). This means that the elements on the same card need not be related, but that elements needed for combinations must actually be spread out carefully over different cards, so it remains possible to form all of the less frequent combinations with cards from one single deck.

At the start of the game, the players are dealt nine cards. Then follow three rounds of play. In each round a student first draws a new card, either from the draw pile or from earlier discards, then scores one combination and finally discards one card. Forming a combination brings a number of points, which is determined by multiplying the point scores for the function and the category (e.g. CO:4 x AJP:3 = 12). If no combination can be formed, three non-combining cards must be 'scored' and ten points are deducted from the player's score.

The limitation to three rounds of play helps keep up the pace of the game, but also puts pressure on the players to focus on all combinations. During the first round, there are generally several options and the highest scoring one can be selected. During the second round, the players have to find a balance between getting a high score in this round and avoiding a point deduction during the last round. During the last round, finally, the players usually have to look for that vital combination-completing card in the discards or else hope for a lucky draw.

After reading a short description of the function and category indications (somewhat more extensive than that in Appendix A), the students learn all they need to learn by playing the game, typically in groups of three or four. They have to analyse all the utterances in their hand and in the available discards in order to form the best-scoring combinations, and they

---

[1] A list of the categories and functions used in this paper can be found in Appendix A.

[2] Card decks are typically printed in sheets of 54 or 55 cards. 54 is also a good number for do-it-yourself construction, as 54 cards can be printed as 6 sheets (A4 or letter) with 9 cards each. For Ling Rummy, a single 54-card deck suffices.

have to check the combinations played by their opponents.[3] If there is a disagreement among the players, they can refer to an accompanying booklet containing the analyses as made during the original annotation of the corpus. If the students have a problem with the analysis found in the booklet, they will have to call on the teacher for arbitration and/or more explanation.

The students are likely to give special attention to the less frequent functions and categories because of their higher scores. More frequent combinations do not really need all that much attention, but are guaranteed to get some anyway when the student wants to avoid the point deduction in the last round and needs to prepare a sure combination. As all functions and categories, as well as most combinations, have to be present in the single deck, their frequencies deviate from those in real text. As a result, the students will not get the right feeling for those frequencies by playing this game. However, some indication is given by the difference in scores. Furthermore, the students do not play this game long enough to develop erroneous intuitions about frequencies and the actual frequencies get sufficient attention in the board game which follows.

3   A Board Game on Syntactic Analysis

The next step in getting to know parsers is the actual complete analysis of whole utterances in the way that a parser is supposed to do it. This necessarily takes some more time, say one to two hours. Also, much more information needs to be presented at the same time, which is not possible in a card game format but acceptable in a board game. The rules can be a bit more complicated as well, but not much, as the focus has to remain on syntax. However, in this particular case, a rule mechanism is needed to force the players to pay attention to each others' analysis activities and not only to their own. This interaction can be achieved by having

players control elements that other players need in their analysis activities.

There are at least two natural models for this type of control. The first is a kind of trading game where grammar rewrites are pictured as trading a constituent for its immediate constituents, with transaction costs (partially) dependent on the likelihood of the rewrite. The players might be able to have monopolies in certain rewrites and would have to be paid by other players wishing to do those rewrites. The exact rules can be adopted from one of the many money, stock or property trading games in existence. The second option is a kind of travel game. The analysis process is then pictured as a journey along a network, e.g. a recursive transition network (RTN; Woods, 1970). Again, players can control parts of the analysis process, by owning sections of the network, which are needed by other players during their journey. Again, there is sufficient inspiration for exact rules, now to be found in one of the many railroad games in existence. I have chosen the second option, the railroad-type game. The main reason is that a trading game tends to lead to too much interaction, typically when players keep spending way too much time on getting better deals. This takes attention away from the actual focus of the game, the analysis, far more than is desired in our setting. Furthermore, the RTN representation is a much more attractive visualization of the parsing process, which is of course very important for a board game. The main disadvantage of the chosen option is that RTN's are hardly ever used in this specific form any more. However, their link to context free grammars should be readily understandable for most students. Also, similar networks are still in use, e.g. in the form of finite state machines.

The details of the resulting game, called the RTN Game, are inspired on the railroad game Box Cars (Erickson and Erickson, 1974, later republished as Rail Baron, Erickson et al., 1977), in which players move their train markers between cities in the United States and can buy historical railroads like the Southern Pacific. In the RTN Game, the map of the United States is replaced by a number of subboards, depicting networks for Adjective Phrase (Figure 2), Adverb Phrase, Noun Phrase, Prepositional Phrase, Sentence (Figure 3)[4] and Utterance, and

---

[3] In the standard game, a scoring player points out the combination and the others merely check if they agree whether the combination is correct. It is also possible to have the other players try to find a combination in the three cards themselves, possibly for a (partial) score. However, this would lead to a much slower game.

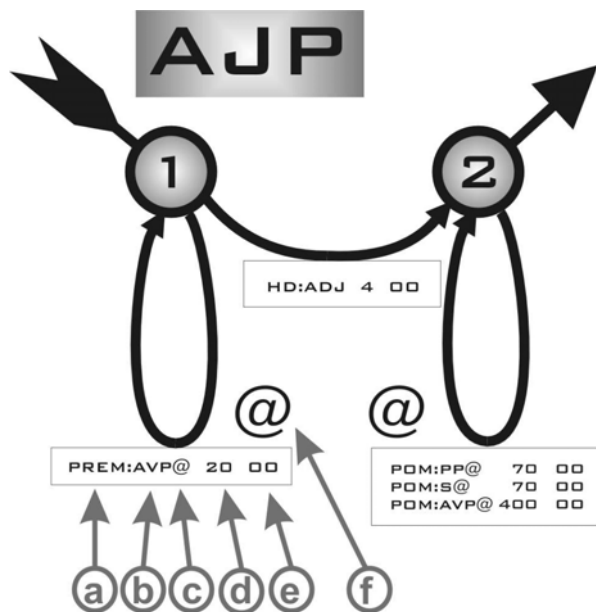
[4] The Sentence subboard is different from the others

Figure 2: The Adjective Phrase subboard in The RTN Game, showing network nodes and arcs marked with a) a syntactic function, b) a syntactic category , c) an @ if the category is non-terminal and hence needs recursion, d) the cost for the arc, e) spaces to mark possession of the arc and f) a space to mark the current analysis position when recursing

players move their pawns through the networks in accordance with the analysis of specific sentences. They can buy network arcs on the board and get paid if anyone uses their arcs. As for Ling Rummy, the optimum number of players is four, but the game will work well with three or five players.

The main activity during the game, then, is moving along the board, which corresponds to analysing specific utterances. The utterances are again taken from the abovementioned corpus sample. However, since the game board should not become too cluttered, the RTN has to be limited in complexity, and the most infrequent constructions are left out. The remaining RTN covers about half of the utterances, several hundred of which are provided on game cards. A

few simple examples can be seen on the Ling Rummy cards in Figure 1. However, there are also more involved utterances, one of the longest being *All that had been proved so_far was that thought could be transferred from one mind to another sometimes, and that the process could be mechanically assisted, at_least as_far_as reception was concerned.*[5] The correct analyses for all utterances (i.e. the analyses selected by the linguist who annotated the sample) are provided in an accompanying booklet which can be consulted if problems arise.

At the start of the game, the players each get three utterance cards from which they can select one to analyse. The analysis consists of moving a pawn (at a die-roll determined speed) along the nodes of the RTN in accordance to the structure of the utterance. Whenever the pawn encounters an arc marked with a recursion sign (@), there is a jump to another network. The current position of the pawn is marked on the larger @ next to the arc and the pawn is then placed at the start of the corresponding network subboard. After the recursion is finished, the pawn returns to the marked position. When moving along an arc, the players have to pay for the use of that arc, e.g. in the AJP network (Figure 2) a premodifying AVP costs 20 (and leads to a detour along the AVP network). The cost for each arc is determined by its frequency of use in the corpus sample; higher cost corresponds to lower frequency. After completion of the analysis of an utterance, the player receives about one and a half times the total cost of that utterance, so that player capital grows throughout the game. Also, after receiving payment, a player is allowed to buy an arc (which has to be paid to the 'bank' and always costs 20), and from that moment on receives the payment from anyone using that arc. Immediately after buying a new arc, the player again draws three utterance cards and selects one of them. The game ends after a fixed amount of time, the winner being the player with the highest amount of money.

In the RTN Game, the players' choices consist of buying the right arcs and selecting the right utterances to analyse. Both types of

---

in that it has three exit paths, one for intransitive sentence patterns (*John sleeps*), one for transitive sentence patterns (*John sees Peter*) and one for intensive sentence patterns (*John is ill*). These three cannot be spread out over several boards because there are arcs for coordinated structures which jump back from the separate parts to the common part.

[5] Compound words have been connected with underscores, e.g. *as_far_as*, and have to be treated as if they are single words.

Figure 3: See separate figure at end of file

Figure 3: The Sentence subboard in The RTN Game (cf. Figure 2 and Footnote 4).

choices force the desired involvement with other players' activities. If another player's current utterance route contains a high-value arc, buying that arc will bring an instant return, and it is therefore useful to (partly) analyse the other players' utterances. If no short-term gain is identified, an arc has to be selected which has good long term prospects. Thinking about these prospects brings insight into grammatical probabilities, as the costs of the arcs depend on the frequencies of occurrence. For utterance selection, the aspects to be considered are the ownership of the needed arcs and the time it will take to traverse the network, i.e. how long it takes before something new can be bought. Again, analysis skills and probabilistic reasoning are honed. Even more than in Ling Rummy, there may be disagreements about analyses, which can either be resolved by referring to the accompanying booklet with 'gold standard' analyses or by discussion with the teacher.

Throughout the game, the students experience what a parser does and in which terms it 'sees' the analysis process. They do not yet experience what the parser cannot do, as all the utterances in the game can be parsed with the RTN on the board. However, this experience can now be provided with a few simple questions, such as "Which utterances in text X cannot be parsed with this RTN?" or "How would the RTN have to be extended to parse utterance Y?". Their experience with the existing RTN should form a sufficient basis for a discussion of such subjects.

4   A Role-Playing Game on Parsing Algorithms

When playing the RTN game, the students have total information about the sentence, as well as their linguistic and world knowledge, and should therefore be able to choose the right path through the network immediately, even though the utterance may be globally or locally ambiguous. They may or may not realize that a parser has more limited knowledge and hence

You are the AJP (adjective phrase) component of the parser. You will be called upon to give information about the presence and extent of AJP's in an input utterance.

You know that an AJP is composed of (in this order)
1. zero or more premodifiers, each realized by an AVP (P=20%)
2. a head, realized by an adjective
3. zero or more postmodifiers, each realized by either a PP (P=5%), an S (P=5%) or an AVP (P=1%)

However, you cannot see the input itself. If you need to know if any potential constituents of an AJP exist at specific positions in the input, you will have to ask your fellow phrase structure components about the existence of PP's, S's or AVP's, or the lexical-morphological component about what kind of word is present at a certain position in the input.

Apart from the knowledge above, and the communication channels to the other components, you have access to processing power (your brain) and your own bit of memory (paper, blackboard or whiteboard).

Figure 4: Instructions for the AJP player in A&A

more trouble picking a route. This realization can again be induced with a few direct questions, such as "In utterance X, how can the parser know whether to pick arc Y or Z at point P?". However, if there is sufficient time, it may be more useful to let the students each take the role of a parser component and get a wider experience. This can be done in a roleplaying game called Analyses and Ambiguities (A&A).

In A&A, each player plays a component of a parser, either one of the constituent-based components that are also present in the RTN Game or a lower-level component like a tokenizer or a lexical-morphological analyzer. Each component has only limited knowledge about the world and a limited access to the input. The AJP component, e.g., will know that an AJP may start with premodifying AVP's, but has to call on the AVP component to find out if there are actually AVP's present at the current location. Also, it has to call on the lexical-morphological component to find out if the current word is an adjective and can hence be used as a adjective phrase head. After gaining information about the accessibility of the different arcs, it may have to choose between various competing routes. Finally, it has to keep an administration in order to be able to keep track of its various instantiations in case of recursion, and possibly for purposes of

backtracking. By experiencing the process at this level, the students learn how the individual components have to do their work.

In addition, the controlling intelligence of the parser is played by all the players together. They decide as a group how to use the component knowledge to perform the overall parsing task. If necessary, they can create a central administration area (e.g. on a blackboard) to control the process as a whole. If there are students without roles, the group might assign one of them to take on the role of a new component, such as a separate recursion administrator or an analysis tree builder.

The group can experiment with various strategies like top-down or bottom-up parsing, look-ahead, parallel parsing, shared forests and probabilistic ordering. At the start, they should be allowed to come up with these strategies themselves, but it is likely that some hints from the teacher will be needed at some point. Alternatively, different teams can be instructed to investigate different strategies, e.g. top-down versus bottom-up, and given time to develop a system using the given strategies. After each team has finished, they can demonstrate their resulting system to the whole group and the relative merit of the systems can be discussed.

The natural group size for the use of A&A is the number of components in the system, i.e. eight if playing the six RTN's plus a tokenizer and a lexical-morphological analyzer. With more students there is a choice between splitting the component parts into subparts, or having the additional students only take part in the group discussion. However, care must be taken that all students are actively involved in the game, and experience shows that attention tends to wander if the group is larger than ten to fifteen students.[6]

---

[6] A&A itself has not been used with students as yet, but the same teaching technique is used at the computer science department, where it is well-appreciated by the students (Wupper, 1999). Here, the technique is known as *technodrama,* indicating parallels with the *psychodrama* used in psychology. More information, including video's of students in action, although only in Dutch, can be found at www.cs.kun.nl/ita/onderwijs/onderwijsvormen/ technodrama/uitleg.html.

5   Conclusion

In any course where the students are to be taught about parsing by personal experience, the three games described in this paper can be used directly as a teaching aid.[7] For lower level courses, Ling Rummy and The RTN Game can be completed with a few simple extra assignments to give a good impression of what a parser does, what goes on inside and how an underlying grammar should be constructed, all within two to three hours. For more computationally minded students, the algorithmic complexities of the parsing process can be learned through A&A, probably taking another hour or two. In both cases, neither previous skills nor access to computing facilities are necessary.

It should be noted that all three games are as yet at the playtesting stage. We plan to use Ling Rummy and The RTN Game for the first time in an actual classroom setting during a first-year linguistics course later this year. The most important lesson I hope to learn then is how university students react to being asked to play 'games'. I expect the majority to react well, but some might well scoff at such 'childish' activities. For these students, the presentation may have to be altered. A minimal alteration is a mere change in terminology. The word *games* can be weakened, e.g. into *game-like activities*, or avoided altogether, leading to terms like *simulation* or *technodrama* (cf. Footnote 6). A step further would be to remove all game elements, like scoring. The same game boards and cards can also be used for straightforward simulations and/or exercises. However, I expect the removal of the game elements to have a detrimental effect on the average student's involvement.

If the games are received well, the road is open to further teaching games. The description of the games in this paper is therefore meant as more than just an introduction to these specific games. It is also intended as a demonstration of how to translate your teaching goals into games, and how conditions on the teaching goals and situation should influence the game format and rules. The game should after all not become an end in itself, but should clearly be a vehicle for teaching the appropriate lessons. In some cases it may be impossible to create a game that is both playable and teaches the desired lessons, but it is my contention that games can certainly be developed for many more aspects of NLP/CL.

---

[7] The games are currently completely in English (subject matter, game materials and instructions). A version of Ling Rummy for Dutch grammatical analysis is under consideration. All games are freely available for teaching purposes, the only condition being that evaluative feedback is provided. Contact the author if you are interested.

References

A Lady. 1802. *Grammatical Game in Rhyme*. Cheapside, London: Saml. Conder.

Allen, L.E., P. Kugel and J.K. Ross. 1970. *Queries 'n Theories*. Fairfield, IA: WFF 'N PROOF Learning Games Associates.

Allingham, M. 1965. *The Mind Readers*. London: Chatto & Windus.

Erickson, T.F. Jr. and R.S. Erickson. 1974. *Box Cars*. Atlanta, Georgia: Erickson.

Erickson, T.F. Jr., R.S. Erickson, R. Hamblen and M. Uhl. 1977. *Rail Baron*. Baltimore, Maryland: The Avalon Hill Game Co.

Nouza, J., N. Holada and D. Hajek. 1997. An Educational and Experimental Workbench for Visual Processing of Speech Data. In *Proc. Eurospeech'97, Rhodes, Greece, September 1997*, pages 661-664.

Oostdijk, N. 2000. English descriptive linguistics at a cross-roads. *English Studies*, Vol. 81,2:127-141.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Woods, W.A. 1970. Transition network grammars for natural language analysis. *CACM* 13(10): 591-606.

Wupper, H. 1999. Anatomy of Computer Systems - Experiences with a new introductory informatics course. Report CSI-R9914. Nijmegen: CSI.

Appendix A

In the examples and figures in this paper, syntactic categories and functions are often indicated by their abbreviated name. This appendix contains a list of these abbreviations and the corresponding full names.

The various categories are:

Adj	Adjective
Adv	Adverb
AJP	Adjective Phrase
Art	Article
Aux	Auxiliary verb
AVP	Adverb Phrase
Con	Connective (adverb)
Conj	Conjunction
Lv	Lexical verb
N	Noun
NP	Noun Phrase
PP	Prepositional Phrase
S	Sentence
To	Infinitival *to*

The various functions are:

A	Adverbial
AV	Auxiliary verb
CO	Object complement
COOR	Coordinator
CS	Subject complement
HD	Head
MV	Main verb
OD	Direct object
OI	Indirect object
PC	Prepositional complement
POM	Postmodifier
PREM	Premodifier
PRTCL	Particle
SU	Subject
SUB	Subordinator

S

COOR:CONJ
400 00

COOR:CONJ
400 00

COOR:CONJ
80 00

INTRANS

COOR:CONJ
400 00

COOR:CONJ
100 00

COOR:CONJ
400 00

TRANS

COOR:CONJ
400 00

COOR:CONJ
400 00

1   2   3   4   5   6

SU:NP@   2 00
SU:<>   12 00
SU:S@   160 00

@

MV:LV 2 00

OI:<>   4 00
OI:NP@   100 00

@

OD:NP@   6 00
OD:S@   16 00

@

CO:<>   4 00
CO:PP@   130 00
CO:AJP@ 150 00
CO:NP@   400 00
CO:S@   400 00

@

SUB:CONJ 14 00
A:AVP@   40 00
A:CON   60 00
A:S@   60 00
A:PP@   220 00
A:NP@   400 00

@

AV:AUX   4 00
PRTCL:TO 22 00
A:AVP@   30 00
A:S@   400 00
A:PP@   400 00
A:NP@   400 00

@

A:PP@   6 00
A:AVP@   10 00
A:S@   30 00
A:NP@   250 00
A:CON   400 00

@

COOR:CONJ
250 00

INTENS

7

CS:AJP@   10 00
CS:NP@   24 00
CS:PP@   50 00
CS:S@   110 00
CS:AVP@ 350 00

@

A:S@   80 00
A:PP@   100 00
A:AVP@   200 00
A:NP@   400 00

@

A:PP@   30 00
A:AVP@   30 00
A:S@   60 00
A:NP@   400 00
A:CON   400 00

@

COOR:CONJ
200 00