Comp349
Reflections on Week 11 Tutorial Submissions

Robert Dale
Robert.Dale@mq.edu.au
Tutorial 11

• Task 1: Date Grammars
• Task 2: Recognizing Years
Caveat

• I have only picked up on some mistakes or problems in these grammars, not all …

• … so don’t assume everything not marked is ok.
Task 1: Date Grammars
Possible Responses #1

- Friday
- Next Friday
- On Friday the 13\textsuperscript{th} -- always spell out in words: ... the thirteenth
- Day after tomorrow -- not natural English: the day after tomorrow
- This coming Friday
- After two days -- not natural English
- Two days from now
- In two days
- In a couple of days -- not a likely response given the ambiguity of ‘couple’
- After 48 hours -- not a likely response
Task 1: Date Grammars
Possible Responses #2

• This Friday.
• Two days later. -- not natural English: ‘two days from now’ more likely
• 19th October.
• I want to Travel on Friday this week.
• Oh, Friday.
• The day after tomorrow.
• After Thursday, should be Friday. -- a little unnatural
• On Friday, the nineteenth.
• I must go on Friday.
• 19th This month. -- not natural English → ‘nineteenth of this month’
Task 1: Date Grammars
Possible Responses #3

- Friday 26th
- Friday 26th October
- Friday 26th October 2007
- Friday 26th of October
- Friday 26th of October 2007
- 26th of the 10th
- 26th of the 10th 2007

Always spell out numbers in words; otherwise you can’t predict the actual behaviour of the speech recognizer, and you may delude yourself into thinking it will recognize things it can’t

Probably not very likely responses
Task 1: Date Grammars
Possible Responses #4

• Friday.
• I’d like to travel on Friday
• I want to travel on Friday
• The twenty-sixth of October
• October twenty-sixth
• The twenty-sixth
• In two days
• In two days time
• Two days from now
• The day after tomorrow
Task 1: Date Grammars
Possible Responses #5

- Friday.
- Friday please.
- This Friday.
- On Friday.
- This coming Friday.
- This week Friday.
- Two days from now.
- In two days.
- On the nineteenth.
- I'd like to go | travel | leave | fly on Friday.
Task 1: Date Grammars
Initial Grammars #1

DATE → RELATIVE-DATE | ABSOLUTE-DATE
RELATIVE-DATE → today | tomorrow | RELATIVE-NP | RELATIVE-PP
ABSOLUTE-DATE → DATE-NP | DATE-PP
RELATIVE-NP → the day after tomorrow | two days later
RELATIVE-PP → ([Oh] DAY-NAME) | ([On] DAY_NAME [the] CARDINALDAYNUMBER )
| ( [I must go on] DAY_NAME) | ([this] DAY-NAME)
| (CARDINALDAYNUMBER [this month])
| ([after] DAY-NAME [should be] DAY-NAME)
| ([I want to travel] on DAY-NAME [this week])
DATE-NP → (MONTH CARDINALDAYNUMBER) | ([the] ORDINALDAYNUMBER [of] MONTH)
DATE-PP → on DATE-NP
MONTH → January | February | March | …
CARDINALDAYNUMBER → one | two | … | thirty-one
ORDINALDAYNUMBER → first | second | … | thirty-first

Why only allow this in some cases? And it’s not a preposition.

Generalise to other number of days.

Why only allow this in some cases?

There needs to be an ‘of’ in here.
Task 1: Date Grammars

Initial Grammars #2

DATE -> day after tomorrow | DAY-MONTH-NP | MONTH-DAY-NP | WEEKDAY-NP | NUMBER-OF-DAYS-NP | NUMBER-OF-HOURS-NP
DAY-MONTH-NP -> (CARDINALDAYNUMBER | ([the] ORDINALDAYNUMBER [of])) MONTH
MONTH-DAY-NP -> MONTH (CARDINALDAYNUMBER | ORDINALDAYNUMBER)
WEEKDAY-NP > [WEEKDAY-START] WEEKDAY [DAY-MONTH-NP | MONTH-DAY-NP] [WEEKDAY-END]
NUMBER-OF-DAYS-NP -> [after | in] | ([CARDINALDAYNUMBER | a couple of ] days [from now]
NUMBER-OF-HOURS-NP -> [after | in] NUMHOURS hours
WEEKDAY-START -> on | next | [this] coming -- ‘coming’ unlikely without ‘this’
WEEKDAY-END -> [this] week
WEEKDAY -> Monday | Tuesday | … | Sunday
MONTH -> January | February | … | December

First rule could be factorised more: eg absolute vs relative.
Task 1: Date Grammars
Initial Grammars #3

DATE → [sometime] (ABSOLUTEDATE | RELATIVEDATE | COMBINATIONDATE)
COMBINATIONDATE → ([on] RELATIVE-PP ABSOLUTEDATE)
ABSOLUTEDATE → DATE-NP | DATE-PP
RELATIVEDATE → today | tomorrow | RELATIVE-NP | RELATIVE-PP
RELATIVE-NP → [the] day after tomorrow
RELATIVE-PP → (([next] | [this]) Friday) | ([in] two days' time)
DATE-NP → (October twenty-six [YEAR]) | (October twenty-sixth [YEAR])
           | ([the] twenty-sixth [of] October [YEAR]) | (twenty-six October [YEAR]) | (twenty-six ten [YEAR])
DATE-PP → on DATE-NP
YEAR → two thousand and seven

This misinterpretation arose in a few submissions: it doesn't really make sense to have a grammar that only recognizes references to a specific day.
Task 1: Date Grammars
Initial Grammars #4

DEPART -> REQUEST | DATE-NP | RELATIVE
REQUEST -> (I’d like | I want) to travel on Friday
DATE-NP -> ([the] twenty-sixth) | DAYMONTH | MONTHDAY
DAYMONTH -> [the] twenty-sixth of October
MONTHDAY -> October twenty-sixth
RELATIVE -> (in two days [time]) | ([in] two days from now) | (the day after tomorrow)

Why can we only use the request form when we say ‘on Friday’? This part could be factored out to be general to more situations: all the DATE-NP and RELATIVE formulations would also be valid here. This failure to locate things at the most general point was another common mistake.
Task 1: Date Grammars
Initial Grammars #5

\[
<\text{friday}\_\text{expr}> = [\text{i'd like to } <\text{verb}>] \\
(<\text{day}\_\text{of}\_\text{week}\_\text{expr}> | <\text{days}\_\text{from}\_\text{now}\_\text{expr}> | <\text{day}\_\text{of}\_\text{month}>); \\
\]

\[
<\text{day}\_\text{of}\_\text{week}\_\text{expr}> = [<\text{specifier}>] \text{ friday } [\text{please}]; \\
<\text{specifier}> = (\text{on} | \text{this } [\text{coming} | \text{week}]); \text{ ‘this week Friday’ sounds odd to me} \\
<\text{verb}> = (\text{go} | \text{travel} | \text{leave} | \text{fly}); \\
\]

\[
<\text{days}\_\text{from}\_\text{now}\_\text{expr}> = (\text{two days from now} | \text{in two days}); \\
<\text{day}\_\text{of}\_\text{month}> = \text{on the nineteenth}; \\
\]

\[
<\text{day}\_\text{of}\_\text{week}> = \\
(\text{monday} | \text{tuesday} | \text{wednesday} | \text{thursday} | \text{friday} | \text{saturday} | \text{sunday}); \\
\]

A neat solution, although it also has some of the problems we’ve seen in earlier ones. Note that it’s ok to use another formalism (here, JGSF) as long as you make very clear that’s what you are doing.
Task 1: Date Grammars
Initial Grammars #5

Friday-DATE → Friday | Friday-NP | Friday-PP
Friday-NP → the day after tomorrow
Friday-NP → this [coming] Friday [if possible]
Friday-NP → two (days | day) from now
Friday-PP → on Friday [this week] [please]
Friday-PP → in two days [time]

It's ok to have separate rules rather than using '|' to indicate alternative possibilities on the right hand side, but usually listing the alternatives on the right hand side makes it easier to see what is allowed.

Should be allowed more generally; put in the top level rule.
Tutorial 11

- Task 1: Date Grammars
- Task 2: Recognizing Years
Task 2: Year Grammars
Solution #1

YEAR → this year | last year | next year | YEAR-NP | YEAR-PP
YEAR-NP → YEAR-CENTRY YEAR-YEAR -- use correct spelling!
YEAR-PP → on YEAR | since YEAR | from YEAR to YEAR
YEAR-CENTRY → NINETEEN | TWO
YEAR-YEAR → thousand one | thousand two | ... | thousand ninety-nine | hundred | one | two | three | ... | ninety-nine

Could be factored out into a separate rule for relative years:
REL-YEAR → (this | last | next) year

Not part of the specification.

This overgenerates: ‘nineteen thousand one’ is not a valid year.

Are these grammar rules or terminal symbols?
Task 2: Year Grammars
Solution #2

YEARS -> YEAR-LAST-HALF | YEAR-SPLIT | YEAR-FULL

YEAR-FULL> (nineteen hundred | two thousand) [and] [DIGIT | YEAR-LAST-HALF]
YEAR-SPLIT-> (YEAR-START-AS-DIGIT | YEAR-START-AS-TENS) (YEAR-LAST-HALF)
YEAR-LAST-HALF-> ((DIGIT DIGIT) | TWODIGITS | (TENS [DIGIT]))

YEAR-START-AS-TENS -> (nineteen | twenty)
YEAR-START-AS-DIGIT -> (one nine) | (two ZERO)

TENS -> twenty | thirty | forty | fifty | sixty | seventy | eighty | ninety
TWODIGITS -> ten | eleven | twelve | thirteen | fourteen | fifteen | sixteen | seventeen
| eighteen | nineteen
DIGIT -> ZERO | one | two | three | four | five | six | seven | eight | nine
ZERO -> o | zero

Maybe a little unlikely in a spoken grammar.

Overgenerates a bit: allows, eg, ‘twenty zero’
Task 2: Year Grammars
Solution #3

YEAR → DIGITS | THOUSANDNUMBER | CENTURYNUMBER
THOUSANDNUMBER → THOUSAND | (THOUSAND [and] NUMBER)
THOUSAND → two thousand
CENTURYNUMBER → CENTURIES | (CENTURIES [and] NUMBER)
CENTURIES → (nineteen | twenty) [hundred]
NUMBER → TEENS |
   (TENS [ONES] |
     ((ONES | ZEROS) (ONES | ZEROS)))
TEENS → eleven | twelve | ... | nineteen
TENS → twenty | thirty | ... | ninety
DIGITS → (one nine | two NAUGHTS ) (ONES | NAUGHTS ) (ONES | NAUGHTS)
ONES → one | two | ... | nine
NAUGHTS → ZEROS | naught
ZEROS → zero | oh

Neat solution to centuries.

Maybe a little unlikely in a spoken grammar.
Task 2: Year Grammars
Solution #4

YEAR -> THOUSANDSFIRST | HUNDREDSFIRST
THOUSANDSFIRST -> THOUSANDS [HUNDREDS] [and TENS]
HUNDREDSFIRST -> ((nineteen | twenty) hundred [and TENS]) |
    ( (nineteen | twenty) (zero | oh) ONES) |
THOUSANDS -> [one | two] thousand
HUNDREDS -> nine hundred
TENS -> ONES | TWODIGIT
ONES -> one | two | three | ... | nine
TWODIGIT -> ten | eleven | ... | ninety-eight | ninety-nine

Overgenerates a bit: allows, eg, ‘two thousand nine hundred …’

Rule names a bit counter-intuitive.

Overgenerates.
Task 2: Year Grammars
Solution #5

FirstTwo --> nineteen hundred | two thousand
FirstTwo2 --> nineteen [hundred] | two thousand
LastDigit --> one | two | three | ............ | nine
LastTwoDigit --> ten | eleven | twelve | ............ | nintey-nine
Year-PP --> FirstTwo | FirstTwo and LastDigit | FirstTwo2 [and] LastTwoDigit
DigitForm --> LastDigit LastDigit LastDigit LastDigit
Grammar --> Year-PP | DigitForm

Second rule covers everything the first rule covers, so the first rule is redundant.

More conventional form is to use three dots: ‘…’ (called ‘ellipsis’).

These are not prepositional phrases so this is an odd choice of rule name.
Task 2: Year Grammars
Solution #6

\[
\text{<year\_range>} = (\text{<twentieth\_century\_expr> | <twenty\_first\_century\_expr>});
\]
\[
\text{<twenty\_first\_century\_expr>} = (\text{<two\_thou\_expr> | <twenty\_expr>});
\]
\[
\text{<twenty\_expr>} = \text{twenty} (\text{<teens\_expr> | <twenties\_to\_nineties\_expr>});
\]
\[
\text{<two\_thou\_expr>} = \text{two thousand} [[[\text{and}]} (\text{<single\_digit\_year> | <teens\_expr> | <twenties\_to\_nineties\_expr>})];
\]
\[
\text{<twentieth\_century\_expr>} = \text{nineteen}
\quad (\text{<less\_than\_ten\_expr> | <teens\_expr> | <twenties\_to\_nineties\_expr>});
\]
\[
\text{<less\_than\_ten\_expr>} = (\text{hundred} | \text{hundred and} \text{<single\_digit\_year>} | \text{oh} \text{<single\_digit\_year>});
\]
\[
\text{<teens\_expr>} =
\quad (\text{ten} | \text{eleven} | \text{twelve} | \text{thirteen} | \text{fourteen} | \text{fifteen} | \text{sixteen} | \text{seventeen} | \text{eighteen} | \text{nineteen});
\]
\[
\text{<twenties\_to\_nineties\_expr>} = \text{<decades\_above\_ten>} [\text{<single\_digit\_year>}];
\]
\[
\text{<decades\_above\_ten>} = (\text{twenty} | \text{thirty} | \text{forty} | \text{fifty} | \text{sixty} | \text{seventy} | \text{eighty} | \text{ninety});
\]
\[
\text{<single\_digit\_year>} = (\text{one} | \text{two} | \text{three} | \text{four} | \text{five} | \text{six} | \text{seven} | \text{eight} | \text{nine});
\]
Task 2: Year Grammars
Solution #7

YEAR → nineteen hundred
YEAR → (nineteen | twenty) oh DIGIT
YEAR → (nineteen | twenty) TWODIGIT
YEAR → two thousand
YEAR → two thousand and DIGIT
DIGIT → one | two | three | ... | nine
TWODIGIT → ten | eleven | twelve | thirteen | ... | (twenty DIGIT) | (thirty DIGIT) | ... | (ninety DIGIT)

Could be factorised better.