Python code for Artificial Intelligence: Foundations of Computational Agents

David L. Poole and Alan K. Mackworth

Version 0.9.4 of March 2, 2022.
# Contents

## 1 Python for Artificial Intelligence

1.1 Why Python? ................................................. 9
1.2 Getting Python ........................................... 9
1.3 Running Python .......................................... 10
1.4 Pitfalls .................................................... 11
1.5 Features of Python ....................................... 11
  1.5.1 Lists, Tuples, Sets, Dictionaries and Comprehensions  11
  1.5.2 Functions as first-class objects ........................ 12
  1.5.3 Generators and Coroutines ............................ 14
1.6 Useful Libraries .......................................... 15
  1.6.1 Timing Code .......................................... 15
  1.6.2 Plotting: Matplotlib .................................. 16
1.7 Utilities .................................................. 16
  1.7.1 Display ............................................... 16
  1.7.2 Argmax ................................................ 18
  1.7.3 Probability ........................................... 19
  1.7.4 Dictionary Union ..................................... 19
1.8 Testing Code ............................................. 19

## 2 Agents and Control

2.1 Representing Agents and Environments ..................... 21
2.2 Paper buying agent and environment ....................... 22
  2.2.1 The Environment .................................... 22
  2.2.2 The Agent ............................................ 24
### Contents

5.5 Assumables .......................................................... 97

6 Planning with Certainty .......................... 101
   6.1 Representing Actions and Planning Problems ........... 101
      6.1.1 Robot Delivery Domain ................................ 102
      6.1.2 Blocks World ........................................... 104
   6.2 Forward Planning ........................................ 106
      6.2.1 Defining Heuristics for a Planner .................. 109
   6.3 Regression Planning .................................... 111
      6.3.1 Defining Heuristics for a Regression Planner .... 113
   6.4 Planning as a CSP ...................................... 114
   6.5 Partial-Order Planning .................................. 117

7 Supervised Machine Learning .................. 125
   7.1 Representations of Data and Predictions .............. 126
      7.1.1 Creating Boolean Conditions from Features ....... 129
      7.1.2 Evaluating Predictions ................................ 131
      7.1.3 Creating Test and Training Sets .................. 132
      7.1.4 Importing Data From File ............................ 133
      7.1.5 Augmented Features .................................. 136
   7.2 Generic Learner Interface ............................. 138
   7.3 Learning With No Input Features ...................... 139
      7.3.1 Evaluation ............................................ 141
   7.4 Decision Tree Learning ............................... 143
   7.5 Cross Validation and Parameter Tuning ............... 147
   7.6 Linear Regression and Classification ................ 150
      7.6.1 Batched Stochastic Gradient Descent ............ 156
   7.7 Boosting ................................................. 157
      7.7.1 Gradient Tree Boosting .............................. 160

8 Neural Networks and Deep Learning .......... 163
   8.1 Layers .................................................... 163
   8.2 Feedforward Networks ................................. 166
   8.3 Improved Optimization .................................. 168
      8.3.1 Momentum ............................................. 168
      8.3.2 RMS-Prop ............................................. 169
      8.3.3 Examples ............................................. 170

9 Reasoning Under Uncertainty ............... 175
   9.1 Representing Probabilistic Models .................... 175
   9.2 Representing Factors ..................................... 176
   9.3 Conditional Probability Distributions ............... 177
      9.3.1 Logistic Regression .................................. 178
      9.3.2 Noisy-or .............................................. 179
      9.3.3 Tabular Factors ...................................... 179

http://aipython.org Version 0.9.4 March 2, 2022
Contents

13.1 Representing Agents and Environments .............................................. 267
  13.1.1 Simulating an environment from an MDP ................................. 268
  13.1.2 Simple Game .................................................................. 269
  13.1.3 Evaluation and Plotting ...................................................... 271
13.2 Q Learning .................................................................................. 273
  13.2.1 Testing Q-learning .............................................................. 275
13.3 Q-learning with Experience Replay ................................................. 276
13.4 Model-based Reinforcement Learner ................................................. 278
13.5 Reinforcement Learning with Features ............................................. 281
  13.5.1 Representing Features .......................................................... 281
  13.5.2 Feature-based RL learner ....................................................... 284
  13.5.3 Experience Replay ............................................................... 287
13.6 Multiagent Learning ...................................................................... 288

14 Relational Learning ........................................................................... 295
  14.1 Collaborative Filtering ................................................................. 295
    14.1.1 Alternative Formulation ....................................................... 298
    14.1.2 Plotting ........................................................................ 298
    14.1.3 Creating Rating Sets ......................................................... 299

15 Version History .................................................................................. 303

Bibliography ......................................................................................... 305

Index ...................................................................................................... 307
1.1 Why Python?

We use Python because Python programs can be close to pseudo-code. It is designed for humans to read.

Python is reasonably efficient. Efficiency is usually not a problem for small examples. If your Python code is not efficient enough, a general procedure to improve it is to find out what is taking most the time, and implement just that part more efficiently in some lower-level language. Most of these lower-level languages interoperate with Python nicely. This will result in much less programming and more efficient code (because you will have more time to optimize) than writing everything in a low-level language. You will not have to do that for the code here if you are using it for course projects.

1.2 Getting Python

You need Python 3 (http://python.org/) and matplotlib (http://matplotlib.org/) that runs with Python 3. This code is not compatible with Python 2 (e.g., with Python 2.7).

Download and install the latest Python 3 release from http://python.org/. This should also install pip3. You can install matplotlib using

```
pip3 install matplotlib
```

in a terminal shell (not in Python). That should “just work”. If not, try using pip instead of pip3.

The command python or python3 should then start the interactive python shell. You can quit Python with a control-D or with quit().
To upgrade matplotlib to the latest version (which you should do if you install a new version of Python) do:

```
pip3 install --upgrade matplotlib
```

We recommend using the enhanced interactive python ipython (http://ipython.org/). To install ipython after you have installed python do:

```
pip3 install ipython
```

### 1.3 Running Python

We assume that everything is done with an interactive Python shell. You can either do this with an IDE, such as IDLE that comes with standard Python distributions, or just running ipython3 (or perhaps just ipython) from a shell.

Here we describe the most simple version that uses no IDE. If you download the zip file, and cd to the “aipython” folder where the .py files are, you should be able to do the following, with user input following:

```
ipython -i searchGeneric.py
```

Testing problem 1:

```
Path found: a --> b --> c --> d --> g
Passed unit test
```

```
In [1]: searcher2 = AStarSearcher(searchProblem.acyclic_delivery_problem)  #A*
```

```
In [2]: searcher2.search()  # find first path
16 paths have been expanded and 5 paths remain in the frontier
Out[2]: o103 --> o109 --> o119 --> o123 --> r123
```

```
In [3]: searcher2.search()  # find next path
21 paths have been expanded and 6 paths remain in the frontier
Out[3]: o103 --> b3 --> b4 --> o109 --> o119 --> o123 --> r123
```

```
In [4]: searcher2.search()  # find next path
28 paths have been expanded and 5 paths remain in the frontier
Out[4]: o103 --> b3 --> b1 --> b2 --> b4 --> o109 --> o119 --> o123 --> r123
```

```
In [5]: searcher2.search()  # find next path
No (more) solutions. Total of 33 paths expanded.
```

http://aipython.org  Version 0.9.4  March 2, 2022
1.4 Pitfalls

It is important to know when side effects occur. Often AI programs consider what would happen or what may have happened. In many such cases, we don’t want side effects. When an agent acts in the world, side effects are appropriate.

In Python, you need to be careful to understand side effects. For example, the inexpensive function to add an element to a list, namely `append`, changes the list. In a functional language like Haskell or Lisp, adding a new element to a list, without changing the original list, is a cheap operation. For example if \( x \) is a list containing \( n \) elements, adding an extra element to the list in Python (using `append`) is fast, but it has the side effect of changing the list \( x \). To construct a new list that contains the elements of \( x \) plus a new element, without changing the value of \( x \), entails copying the list, or using a different representation for lists. In the searching code, we will use a different representation for lists for this reason.

1.5 Features of Python

1.5.1 Lists, Tuples, Sets, Dictionaries and Comprehensions

We make extensive uses of lists, tuples, sets and dictionaries (dicts). See https://docs.python.org/3/library/stdtypes.html

One of the nice features of Python is the use of list comprehensions (and also tuple, set and dictionary comprehensions).

\[(fe \text{ for } e \text{ in } iter \text{ if } cond)\]

enumerates the values \( fe \) for each \( e \) in \( iter \) for which \( cond \) is true. The “if \( cond \)” part is optional, but the “for” and “in” are not optional. Here \( e \) has to be a variable, \( iter \) is an iterator, which can generate a stream of data, such as a list, a set, a range object (to enumerate integers between ranges) or a file. \( cond \)
is an expression that evaluates to either True or False for each $e$, and $fe$ is an expression that will be evaluated for each value of $e$ for which $cond$ returns True.

The result can go in a list or used in another iteration, or can be called directly using $next$. The procedure $next$ takes an iterator returns the next element (advancing the iterator) and raises a $StopIteration$ exception if there is no next element. The following shows a simple example, where user input is prepended with >>>

```python
>>> [e*e for e in range(20) if e%2==0]
[0, 4, 16, 36, 100, 144, 196, 256, 324]
>>> a = (e*e for e in range(20) if e%2==0)
>>> next(a)
0
>>> next(a)
4
>>> next(a)
16
>>> list(a)
[36, 64, 100, 144, 196, 256, 324]
```

Notice how $list(a)$ continued on the enumeration, and got to the end of it.

Comprehensions can also be used for dictionaries. The following code creates an index for list $a$:

```python
>>> a = ['a','f','bar','b','a','aaaaa']
>>> ind = {a[i]:i for i in range(len(a))}
>>> ind
{'a': 4, 'f': 1, 'bar': 2, 'b': 3, 'aaaaa': 5}
```

which means that 'b' is the 3rd element of the list.

The assignment of $ind$ could have also be written as:

```python
>>> ind = {val:i for (i,val) in enumerate(a)}
```

where $enumerate$ returns an iterator of $(index, value)$ pairs.

### 1.5.2 Functions as first-class objects

Python can create lists and other data structures that contain functions. There is an issue that tricks many newcomers to Python. For a local variable in a function, the function uses the last value of the variable when the function is
1.5. Features of Python

called, not the value of the variable when the function was defined (this is called “late binding”). This means if you want to use the value a variable has when the function is created, you need to save the current value of that variable. Whereas Python uses “late binding” by default, the alternative that newcomers often expect is “early binding”, where a function uses the value a variable had when the function was defined, can be easily implemented.

Consider the following programs designed to create a list of 5 functions, where the $i$th function in the list is meant to add $i$ to its argument.\[1\]

```python
fun_list1 = []
for i in range(5):
    def fun1(e):
        return e+i
    fun_list1.append(fun1)

fun_list2 = []
for i in range(5):
    def fun2(e, iv=i):
        return e+iv
    fun_list2.append(fun2)

fun_list3 = [lambda e: e+i for i in range(5)]

fun_list4 = [lambda e, iv=i: e+iv for i in range(5)]
```

Try to predict, and then test to see the output, of the output of the following calls, remembering that the function uses the latest value of any variable that is not bound in the function call:

```python
i=56
# in Shell do
## ipython -i pythonDemo.py
# Try these (copy text after the comment symbol and paste in the Python prompt):
# print([f(10) for f in fun_list1])
# print([f(10) for f in fun_list2])
# print([f(10) for f in fun_list3])
# print([f(10) for f in fun_list4])
```

In the first for-loop, the function $fun$ uses $i$, whose value is the last value it was assigned. In the second loop, the function $fun2$ uses $iv$. There is a separate $iv$ variable for each function, and its value is the value of $i$ when the function was defined. Thus $fun1$ uses late binding, and $fun2$ uses early binding. $fun_list3$

\[1\]Numbered lines are Python code available in the code-directory, aipython. The name of the file is given in the gray text above the listing. The numbers correspond to the line numbers in that file.

http://aipython.org Version 0.9.4 March 2, 2022
and `fun_list4` are equivalent to the first two (except `fun_list4` uses a different `i` variable).

One of the advantages of using the embedded definitions (as in `fun1` and `fun2` above) over the lambda is that is it possible to add a `__doc__` string, which is the standard for documenting functions in Python, to the embedded definitions.

### 1.5.3 Generators and Coroutines

Python has generators which can be used for a form of coroutines.

The `yield` command returns a value that is obtained with `next`. It is typically used to enumerate the values for a `for` loop or in generators. (The `yield` command can also be used for coroutines, but we only use it for generators in AIPython.)

A version of the built-in `range`, with 2 or 3 arguments (and positive steps) can be implemented as:

```python
def myrange(start, stop, step=1):
    """enumerates the values from start in steps of size step that are
    less than stop."
    assert step>0, "only positive steps implemented in myrange"
    i = start
    while i<stop:
        yield i
        i += step

print("list(myrange(2,30,3)): ",list(myrange(2,30,3)))
```

Note that the built-in `range` is unconventional in how it handles a single argument, as the single argument acts as the second argument of the function. Note also that the built-in `range` also allows for indexing (e.g., `range(2,30,3)[2]` returns 8), which the above implementation does not. However `myrange` also works for floats, which the built-in `range` does not.

**Exercise 1.1** Implement a version of `myrange` that acts like the built-in version when there is a single argument. (Hint: make the second argument have a default value that can be recognized in the function.)

Yield can be used to generate the same sequence of values as in the example of Section 1.5.1:

```python
def ga(n):
    """generates square of even nonnegative integers less than n""
    for e in range(n):
        if e%2==0:
            yield e*e
a = ga(20)
```

[http://aipython.org](http://aipython.org)  
Version 0.9.4  
March 2, 2022
The sequence of `next(a), and list(a)` gives exactly the same results as the comprehension in Section 1.5.1. It is straightforward to write a version of the built-in `enumerate`. Let’s call it `myenumerate`:

```python
def myenumerate(enum):
    for i in range(len(enum)):
        yield i, enum[i]
```

Exercise 1.2 Write a version of `enumerate` where the only iteration is “for val in enum”. Hint: keep track of the index.

## 1.6 Useful Libraries

### 1.6.1 Timing Code

In order to compare algorithms, we often want to compute how long a program takes; this is called the runtime of the program. The most straightforward way to compute runtime is to use `time.perf_counter()`, as in:

```python
import time
start_time = time.perf_counter()
compute_for_a_while()
end_time = time.perf_counter()
print("Time:", end_time - start_time, "seconds")
```

Note that `time.perf_counter()` measures clock time; so this should be done without user interaction between the calls. On the console, you should do:

```python
start_time = time.perf_counter(); compute_for_a_while(); end_time = time.perf_counter()
```

If this time is very small (say less than 0.2 second), it is probably very inaccurate, and it may be better to run your code many times to get a more accurate count. For this you can use `timeit` ([https://docs.python.org/3/library/timeit.html](https://docs.python.org/3/library/timeit.html)). To use `timeit` to time the call to `foo.bar(aaa)` use:

```python
import timeit
time = timeit.timeit("foo.bar(aaa)",
                    setup="from __main__ import foo,aaa", number=100)
```

The setup is needed so that Python can find the meaning of the names in the string that is called. This returns the number of seconds to execute `foo.bar(aaa)` 100 times. The variable `number` should be set so that the runtime is at least 0.2 seconds.

You should not trust a single measurement as that can be confounded by interference from other processes. `timeit.repeat` can be used for running `timeit` a few (say 3) times. Usually the minimum time is the one to report, but you should be explicit and explain what you are reporting.
1.6.2 Plotting: Matplotlib

The standard plotting for Python is matplotlib ([http://matplotlib.org/](http://matplotlib.org/)). We will use the most basic plotting using the pyplot interface.

Here is a simple example that uses everything we will use.

```python
import matplotlib.pyplot as plt

def myplot(minv, maxv, step, fun1, fun2):
    plt.ion()  # make it interactive
    plt.xlabel("The x axis")
    plt.ylabel("The y axis")
    plt.xscale('linear')  # Makes a 'log' or 'linear' scale
    xvalues = range(minv, maxv, step)
    plt.plot(xvalues, [fun1(x) for x in xvalues],
             label="The first fun")
    plt.plot(xvalues, [fun2(x) for x in xvalues], linestyle='--', color='k',
             label=fun2.__doc__)  # use the doc string of the function
    plt.legend(loc="upper right")  # display the legend

def slin(x):
    """y=2x+7""
    return 2*x+7

def sqfun(x):
    """y=(x-40)^2/10-20""
    return (x-40)**2/10-20

# Try the following:
# from pythonDemo import myplot, slin, sqfun
# import matplotlib.pyplot as plt
# myplot(0, 100, 1, slin, sqfun)
# plt.legend(loc="best")
# import math
# plt.plot([41+40*math.cos(th/10) for th in range(50)],
#           [100+100*math.sin(th/10) for th in range(50)])
# plt.text(40, 100, "ellipse?")
# plt.xscale('log')
```

At the end of the code are some commented-out commands you should try in interactive mode. Cut from the file and paste into Python (and remember to remove the comments symbol and leading space).

1.7 Utilities

1.7.1 Display

In this distribution, to keep things simple and to only use standard Python, we use a text-oriented tracing of the code. A graphical depiction of the code could
override the definition of display (but we leave it as a project).
The method self.display is used to trace the program. Any call

\[
\textit{self.display}(\text{level, to\_print \ldots})
\]

where the level is less than or equal to the value for \textit{max\_display\_level} will be
printed. The \textit{to\_print \ldots} can be anything that is accepted by the built-in \textit{print}
(including any keyword arguments).

The definition of \textit{display} is:

```
class Displayable(object):
    """Class that uses 'display'.
    The amount of detail is controlled by max_display_level
    ""
    max_display_level = 1 # can be overridden in subclasses

    def display(self,level,*args,**nargs):
        """print the arguments if level is less than or equal to the
        current max_display_level.
        level is an integer.
        the other arguments are whatever arguments print can take.
        """
        if level <= self.max_display_level:
            print(*args, **nargs) #if error you are using Python2 not Python3
```

Note that \textit{args} gets a tuple of the positional arguments, and \textit{nargs} gets a dictionary of the keyword arguments). This will not work in Python 2, and will give an error.

Any class that wants to use \textit{display} can be made a subclass of \textit{Displayable}.
To change the maximum display level to say 3, for a class do:

```
Classname.max_display_level = 3
```

which will make calls to \textit{display} in that class print when the value of \textit{level} is less
than-or-equal to 3. The default display level is 1. It can also be changed for
individual objects (the object value overrides the class value).

The value of \textit{max_display_level} by convention is:

0  display nothing
1  display solutions (nothing that happens repeatedly)
2  also display the values as they change (little detail through a loop)
3  also display more details
4 and above  even more detail
In order to implement more sophisticated visualizations of the algorithm, we add a `visualize` “decorator” to the methods to be visualized. The following code ignores the decorator:

```python
def visualize(func):
    """A decorator for algorithms that do interactive visualization.
    Ignored here.
    """
    return func
```

### 1.7.2 Argmax

Python has a built-in `max` function that takes a generator (or a list or set) and returns the maximum value. The `argmax` method returns the index of an element that has the maximum value. If there are multiple elements with the maximum value, one if the indexes to that value is returned at random. `argmaxe` assumes an enumeration; a generator of `(element, value)` pairs, as for example is generated by the built-in `enumerate(list)` for lists or `dict.items()` for dicts.

```python
import random
import math

def argmaxall(gen):
    """gen is a generator of (element,value) pairs, where value is a real.
    argmaxall returns a list of all of the elements with maximal value.
    """
    maxv = -math.inf       # negative infinity
    maxvals = []            # list of maximal elements
    for (e,v) in gen:
        if v>maxv:
            maxvals,maxv = [e], v
        elif v==maxv:
            maxvals.append(e)
    return maxvals

def argmaxe(gen):
    """gen is a generator of (element,value) pairs, where value is a real.
    argmaxe returns an element with maximal value.
    If there are multiple elements with the max value, one is returned at random.
    """
    return random.choice(argmaxall(gen))

def argmax(lst):
    """returns maximum index in a list"
    return argmaxe(enumerate(lst))
```

[http://aipython.org](http://aipython.org)
1.8 Testing Code

```python
# argmax([1,6,3,77,3,55,23])

def argmaxd(dct):
    """returns the arg max of a dictionary dct""
    return argmaxe(dct.items())

# Try:
# argmaxd({2:5,5:9,7:7})
```

**Exercise 1.3** Change argmax to have an optional argument that specifies whether you want the “first”, “last” or a “random” index of the maximum value returned. If you want the first or the last, you don’t need to keep a list of the maximum elements.

1.7.3 Probability

For many of the simulations, we want to make a variable True with some probability. `flip(p)` returns True with probability `p`, and otherwise returns False.

```python
# utilities.py — (continued)

def flip(prob):
    """return true with probability prob""
    return random.random() < prob
```

1.7.4 Dictionary Union

**This is now | in Python 3.9, so will be replaced.**

The function `dict_union(d1, d2)` returns the union of dictionaries `d1` and `d2`. If the values for the keys conflict, the values in `d2` are used. This is similar to `dict(d1, **d2)`, but that only works when the keys of `d2` are strings.

```python
# utilities.py — (continued)

def dict_union(d1,d2):
    """returns a dictionary that contains the keys of d1 and d2.
    The value for each key that is in d2 is the value from d2,
    otherwise it is the value from d1.
    This does not have side effects.
    ""
    d = dict(d1)  # copy d1
    d.update(d2)
    return d
```

1.8 Testing Code

It is important to test code early and test it often. We include a simple form of **unit test**. The value of the current module is in `__name__` and if the module is run at the top-level, it's value is "__main__". See [https://docs.python.org/3/library/__main__.html](https://docs.python.org/3/library/__main__.html)

[http://aipython.org](http://aipython.org)
The following code tests `argmax` and `dict_union`, but only when if `utilities` is loaded in the top-level. If it is loaded in a module the test code is not run.

In your code you should do more substantial testing than we do here, in particular testing the boundary cases.

```python
def test():
    """Test part of utilities""
    assert argmax(enumerate([1,6,55,3,55,23])) in [2,4]
    assert dict_union({1:4, 2:5, 3:4},{5:7, 2:9}) == {1:4, 2:9, 3:4, 5:7}
    print("Passed unit test in utilities")

if __name__ == "__main__":
    test()
```

http://aipython.org

Version 0.9.4

March 2, 2022
Chapter 2

Agents and Control

This implements the controllers described in Chapter 2.

In this version the higher-levels call the lower-levels. A more sophisticated version may have them run concurrently (either as coroutines or in parallel). The higher-levels calling the lower-level works in simulated environments when there is a single agent, and where the lower-level are written to make sure they return (and don’t go on forever), and the higher level doesn’t take too long (as the lower-levels will wait until called again).

2.1 Representing Agents and Environments

An agent observes the world, and carries out actions in the environment, it also has an internal state that it updates. The environment takes in actions of the agents, updates it internal state and returns the percepts.

In this implementation, the state of the agent and the state of the environment are represented using standard Python variables, which are updated as the state changes. The percepts and the actions are represented as variable-value dictionaries.

An agent implements the $\text{go}(n)$ method, where $n$ is an integer. This means that the agent should run for $n$ time steps.

In the following code raise `NotImplementedError()` is a way to specify an abstract method that needs to be overridden in any implemented agent or environment.

```python
import random

class Agent(object):
    def __init__(self, env):
```
2. Agents and Control

```
"""set up the agent"
self.env=env

def go(self,n):
"""acts for n time steps"
raise NotImplementedError("go") # abstract method
```

The environment implements a do(action) method where action is a variable-value dictionary. This returns a percept, which is also a variable-value dictionary. The use of dictionaries allows for structured actions and percepts.

Note that Environment is a subclass of Displayable so that it can use the display method described in Section 1.7.1.

```
from display import Displayable

class Environment(Displayable):
    def initial_percepts(self):
        """returns the initial percepts for the agent"
        raise NotImplementedError("initial_percepts") # abstract method

def do(self,action):
    """does the action in the environment
    returns the next percept"
    raise NotImplementedError("do") # abstract method
```

2.2 Paper buying agent and environment

To run the demo, in folder "aipython", load "agents.py", using e.g.,
ipython -i agents.py, and copy and paste the commented-out commands at the bottom of that file. This requires Python 3 with matplotlib.

This is an implementation of the paper buying example.

2.2.1 The Environment

The environment state is given in terms of the time and the amount of paper in stock. It also remembers the in-stock history and the price history. The percepts are the price and the amount of paper in stock. The action of the agent is the number to buy.

Here we assume that the prices are obtained from the prices list plus a random integer in range [0, max_price_addon) plus a linear "inflation". The agent cannot access the price model; it just observes the prices and the amount in stock.

```
class TP_env(Environment):
http://aipython.org
```

Version 0.9.4 March 2, 2022
2.2. Paper buying agent and environment

```
prices = [234, 234, 234, 234, 255, 255, 275, 275, 211, 211, 211,
234, 234, 234, 234, 199, 199, 275, 275, 234, 234, 234, 234, 255,
255, 260, 260, 265, 265, 265, 270, 270, 255, 255, 260, 260, 265,
265, 265, 150, 150, 265, 265, 270, 270, 255, 255, 260, 260, 265,
265, 265, 265, 270, 270, 255, 255, 260, 260, 265, 265, 265, 265,
270, 270]
max_price_addon = 20 # maximum of random value added to get price

```
2. Agents and Control

2.2.2 The Agent

The agent does not have access to the price model but can only observe the current price and the amount in stock. It has to decide how much to buy.

The belief state of the agent is an estimate of the average price of the paper, and the total amount of money the agent has spent.

```python
class TP_agent(Agent):
    def __init__(self, env):
        self.env = env
        self.spent = 0
        percepts = env.initial_percepts()
        self.ave = self.last_price = percepts['price']
        self.instock = percepts['instock']

    def go(self, n):
        """go for n time steps""
        for i in range(n):
            if self.last_price < 0.9*self.ave and self.instock < 60:
                tobuy = 48
            elif self.instock < 12:
                tobuy = 12
            else:
                tobuy = 0
            self.spent += tobuy*self.last_price
            percepts = env.do({'buy': tobuy})
            self.last_price = percepts['price']
            self.ave = self.ave + (self.last_price - self.ave)*0.05
            self.instock = percepts['instock']
```

Set up an environment and an agent. Uncomment the last lines to run the agent for 90 steps, and determine the average amount spent.

```python
env = TP_env()
ag = TP_agent(env)
#ag.go(90)
#ag.spent/env.time ## average spent per time period
```
2.3 Hierarchical Controller

2.2.3 Plotting

The following plots the price and number in stock history:

```python
import matplotlib.pyplot as plt

class Plot_prices(object):
    """Set up the plot for history of price and number in stock""
    def __init__(self, ag, env):
        self.ag = ag
        self.env = env
        plt.ion()
        plt.xlabel("Time")
        plt.ylabel("Number in stock.
        Price.")

    def plot_run(self):
        """plot history of price and instock""
        num = len(env.stock_history)
        plt.plot(range(num), env.stock_history, label="In stock")
        plt.plot(range(num), env.price_history, label="Price")
        #plt.legend(loc="upper left")
        plt.draw()

# pl = Plot_prices(ag, env)
# ag.go(90); pl.plot_run()
```

2.3 Hierarchical Controller

To run the hierarchical controller, in folder "aipython", load "agentTop.py", using e.g., ipython -i agentTop.py, and copy and paste the commands near the bottom of that file. This requires Python 3 with matplotlib.

In this implementation, each layer, including the top layer, implements the environment class, because each layer is seen as an environment from the layer above.

We arbitrarily divide the environment and the body, so that the environment just defines the walls, and the body includes everything to do with the agent. Note that the named locations are part of the (top-level of the) agent, not part of the environment, although they could have been.

2.3.1 Environment

The environment defines the walls.
2. Agents and Control

2.3.2 Body

The body defines everything about the agent body.

```python
import math
from agents import Environment

class Rob_env(Environment):
    def __init__(self, walls = {}):
        
        self.walls = walls

    def percepts(self):
        return {'rob_x_pos':self.rob_x, 'rob_y_pos':self.rob_y,
                'rob_dir':self.rob_dir, 'whisker':self.whisker(),
                'crashed':self.crashed}

    initial_percepts = percepts # use percept function for initial percepts too

    def do(self, action):
        
        action = {'steer':direction}
```

http://aipython.org  Version 0.9.4  March 2, 2022
2.3. Hierarchical Controller

```python
if self.crashed:
    return self.percepts()

direction = action['steer']
compass_deriv =
    {'left':1,'straight':0,'right':-1}[direction]*self.turning_angle
self.rob_dir = (self.rob_dir + compass_deriv +360)%360 # make in range [0,360)
rob_x_new = self.rob_x + math.cos(self.rob_dir*math.pi/180)
rob_y_new = self.rob_y + math.sin(self.rob_dir*math.pi/180)
path = ((self.rob_x,self.rob_y),(rob_x_new,rob_y_new))

if any(line_segments_intersect(path,wall) for wall in self.env.walls):
    self.crashed = True
    if self.plotting:
        plt.plot([self.rob_x],[self.rob_y],"r*",markersize=20.0)
        plt.draw()
    self.rob_x, self.rob_y = rob_x_new, rob_y_new
    self.history.append((self.rob_x, self.rob_y))
    if self.plotting and not self.crashed:
        plt.plot([self.rob_x],[self.rob_y],"go")
        plt.draw()
        plt.pause(self.sleep_time)
    return self.percepts()
```

This detects if the whisker and the wall intersect. It's value is returned as a percept.

```python
def whisker(self):
    """returns true whenever the whisker sensor intersects with a wall
    ""
    whisk_ang_world = (self.rob_dir-self.whisker_angle)*math.pi/180
    # angle in radians in world coordinates
    wx = self.rob_x + self.whisker_length * math.cos(whisk_ang_world)
    wy = self.rob_y + self.whisker_length * math.sin(whisk_ang_world)
    whisker_line = ((self.rob_x,self.rob_y),(wx,wy))
    hit = any(line_segments_intersect(whisker_line,wall)
              for wall in self.env.walls)
    if hit:
        self.wall_history.append((self.rob_x, self.rob_y))
        if self.plotting:
            plt.plot([self.rob_x],[self.rob_y],"ro")
            plt.draw()
    return hit
```

```python
def line_segments_intersect(linea,lineb):
    """returns true if the line segments, linea and lineb intersect.
    A line segment is represented as a pair of points.
    A point is represented as a (x,y) pair.
```
2. Agents and Control

2.3.3 Middle Layer

The middle layer acts like both a controller (for the environment layer) and an environment for the upper layer. It has to tell the environment how to steer. Thus it calls `env.do()`. It also is told the position to go to and the timeout. Thus it also has to implement `do()`.

```python
from agents import Environment
import math

class Rob_middle_layer(Environment):
    def __init__(self, env):
        self.env = env
        self.percepts = env.initial_percepts()
        self.straight_angle = 11  # angle that is close enough to straight ahead
        self.close_threshold = 2  # distance that is close enough to arrived
        self.close_threshold_squared = self.close_threshold**2  # just compute it once

    def initial_percepts(self):
        return {}

    def do(self, action):
        """action is {'go_to': target_pos, 'timeout': timeout}
        target_pos is (x, y) pair
        timeout is the number of steps to try
        returns {'arrived': True} when arrived is true
        or {'arrived': False} if it reached the timeout
        """
```

"""
This determines how to steer depending on whether the goal is to the right or the left of where the robot is facing.

```python
def steer(self, target_pos):
    if self.percepts['whisker']:
        self.display(3, 'whisker on', self.percepts)
        return "left"
    else:
        gx, gy = target_pos
        rx, ry = self.percepts['rob_x_pos'], self.percepts['rob_y_pos']
        goal_dir = math.acos((gx-rx)/math.sqrt((gx-rx)*(gx-rx) + (gy-ry)*(gy-ry)))*180/math.pi
        if ry>gy:
            goal_dir = -goal_dir
            goal_from_rob = (goal_dir - self.percepts['rob_dir'])+540)%360-180
            assert -180 < goal_from_rob <= 180
            if goal_from_rob > self.straight_angle:
                return "left"
            elif goal_from_rob < -self.straight_angle:
                return "right"
        else:
            return "straight"

def close_enough(self, target_pos):
    gx, gy = target_pos
    rx, ry = self.percepts['rob_x_pos'], self.percepts['rob_y_pos']
    return (gx-rx)**2 + (gy-ry)**2 <= self.close_threshold_squared
```

2.3.4 Top Layer

The top layer treats the middle layer as its environment. Note that the top layer is an environment for us to tell it what to visit.
from agents import Environment

class Rob_top_layer(Environment):
    def __init__(self, middle, timeout=200, locations = {'mail':(-5,10),
       '0103':(50,10), '0109':(100,10), 'storage':(101,51)):
        """middle is the middle layer
        timeout is the number of steps the middle layer goes before giving
        up
        locations is a loc:pos dictionary
        where loc is a named location, and pos is an (x,y) position.
        """
        self.middle = middle
        self.timeout = timeout # number of steps before the middle layer
        should give up
        self.locations = locations

    def do(self, plan):
        """carry out actions.
        actions is of the form {'visit':list_of_locations}
        It visits the locations in turn.
        """
        to_do = plan['visit']
        for loc in to_do:
            position = self.locations[loc]
            arrived = self.middle.do({'go_to':position,
                'timeout':self.timeout})
            self.display(1, "Arrived at", loc, arrived)

2.3.5 Plotting

The following is used to plot the locations, the walls and (eventually) the movement of the robot. It can either plot the movement if the robot as it is going (with the default env.plotting = True), or not plot it as it is going (setting env.plotting = False; in this case the trace can be plotted using pl.plot_run()).

import matplotlib.pyplot as plt

class Plot_env(object):
    def __init__(self, body, top):
        """sets up the plot
        ""
        self.body = body
        plt.ion()
        plt.clf()
        plt.axes().set_aspect('equal')
        for wall in body.env.walls:
            ((x0,y0),(x1,y1)) = wall

http://aipython.org
2.3. Hierarchical Controller

```python
plt.plot([x0, x1], [y0, y1], "-k", linewidth=3)
for loc in top.locations:
    (x, y) = top.locations[loc]
    plt.plot([x], [y], "k<")
    plt.text(x+1.0, y+0.5, loc) # print the label above and to the right
plt.plot([body.rob_x], [body.rob_y], "go")
plt.draw()
```

```python
def plot_run(self):
    """plots the history after the agent has finished.
    This is typically only used if body.plotting==False
    ""
    xs, ys = zip(*self.body.history)
    plt.plot(xs, ys, "go")
    wxs, wys = zip(*self.body.wall_history)
    plt.plot(wxs, wys, "ro")
    plt.draw()
```

The following code plots the agent as it acts in the world:

```python
from agentEnv import Rob_body, Rob_env
env = Rob_env({((20,0),(30,20)), ((70,-5),(70,25))})
body = Rob_body(env)
middle = Rob_middle_layer(body)
top = Rob_top_layer(middle)

# try:
# pl=Plot_env(body,top)
# top.do({'visit':['o109', 'storage', 'o109', 'o103']})
# You can directly control the middle layer:
# middle.do({'go_to':(30,-10), 'timeout':200})
# Can you make it crash?
```

Exercise 2.1: The following code implements a robot trap. Write a controller that can escape the “trap” and get to the goal. See textbook for hints.

```python
# Robot Trap for which the current controller cannot escape:
trap_env = Rob_env({((10,-21),(10,0)), ((10,10),(10,31)),
                    ((30,-21),(30,0)), ((30,10),(30,20)), ((50,-21),(50,31)),
                    ((10,0),(30,0)), ((10,10),(30,10)), ((10,31),(50,31)))
trap_body = Rob_body(trap_env, init_pos=(-1,0,90))
trap_middle = Rob_middle_layer(trap_body)
trap_top = Rob_top_layer(trap_middle, locations={'goal':(71,0)})
# Robot trap exercise:
```

[http://aipython.org](http://aipython.org)
# pl=Plot_env(trap_body,trap_top)
# trap_top.do({'visit':['goal']})
Chapter 3

Searching for Solutions

3.1 Representing Search Problems

A search problem consists of:

• a start node

• a neighbors function that given a node, returns an enumeration of the arcs from the node

• a specification of a goal in terms of a Boolean function that takes a node and returns true if the node is a goal

• a (optional) heuristic function that, given a node, returns a non-negative real number. The heuristic function defaults to zero.

As far as the searcher is concerned a node can be anything. If multiple-path pruning is used, a node must be hashable. In the simple examples, it is a string, but in more complicated examples (in later chapters) it can be a tuple, a frozen set, or a Python object.

In the following code raise NotImplementedError() is a way to specify that this is an abstract method that needs to be overridden to define an actual search problem.
The methods must be overridden to define a search problem."

```python
def start_node(self):
    """returns start node"
    raise NotImplementedError("start_node") # abstract method

def is_goal(self, node):
    """is True if node is a goal"
    raise NotImplementedError("is_goal") # abstract method

def neighbors(self, node):
    """returns a list of the arcs for the neighbors of node"
    raise NotImplementedError("neighbors") # abstract method

def heuristic(self, n):
    """Gives the heuristic value of node n.
    Returns 0 if not overridden."
    return 0
```

The neighbors is a list of arcs. A (directed) arc consists of a `from_node` node and a `to_node` node. The arc is the pair ⟨`from_node, to_node⟩, but can also contain a non-negative cost (which defaults to 1) and can be labeled with an action.

```python
class Arc(object):
    """An arc has a from_node and a to_node node and a (non-negative)
cost"
    def __init__(self, from_node, to_node, cost=1, action=None):
        assert cost >= 0, ("Cost cannot be negative for"
            + str(from_node) + "->" + str(to_node), cost:
            + str(cost))
        self.from_node = from_node
        self.to_node = to_node
        self.action = action
        self.cost = cost

def __repr__(self):
    """string representation of an arc"
    if self.action:
        return str(self.from_node) + "--" + str(self.action) + "-->" + str(self.to_node)
    else:
        return str(self.from_node) + "--" + str(self.to_node)
```

### 3.1.1 Explicit Representation of Search Graph

The first representation of a search problem is from an explicit graph (as opposed to one that is generated as needed).

An explicit graph consists of

[http://aipython.org](http://aipython.org)
3.1. Representing Search Problems

- a list or set of nodes
- a list or set of arcs
- a start node
- a list or set of goal nodes
- (optionally) a dictionary that maps a node to a heuristic value for that node

To define a search problem, we need to define the start node, the goal predicate, the neighbors function and the heuristic function.

class Search_problem_from_explicit_graph(Search_problem):
    """A search problem consists of:
    * a list or set of nodes
    * a list or set of arcs
    * a start node
    * a list or set of goal nodes
    * a dictionary that maps each node into its heuristic value.
    * a dictionary that maps each node into its (x,y) position
    """

def __init__(self, nodes, arcs, start=None, goals=set(), hmap={}, positions={}):
    self.neighs = {}
    self.nodes = nodes
    for node in nodes:
        self.neighs[node]=[]
    self.arcs = arcs
    for arc in arcs:
        self.neighs[arc.from_node].append(arc)
    self.start = start
    self.goals = goals
    self.hmap = hmap
    self.positions = positions

def start_node(self):
    """returns start node"""
    return self.start

def is_goal(self,node):
    """is True if node is a goal"""
    return node in self.goals

def neighbors(self,node):
    """returns the neighbors of node"""
    return self.neighs[node]
3. Searching for Solutions

```python
def heuristic(self,node):
    """Gives the heuristic value of node n. Returns 0 if not overridden in the hmap."""
    if node in self.hmap:
        return self.hmap[node]
    else:
        return 0

def __repr__(self):
    """returns a string representation of the search problem""
    res=""
    for arc in self.arcs:
        res += str(arc)+". "
    return res
```

The following is used for the depth-first search implementation below.

```python
def neighbor_nodes(self,node):
    """returns an iterator over the neighbors of node""
    return (path.to_node for path in self.neighs[node])
```

3.1.2 Paths

A searcher will return a path from the start node to a goal node. A Python list is not a suitable representation for a path, as many search algorithms consider multiple paths at once, and these paths should share initial parts of the path. If we wanted to do this with Python lists, we would need to keep copying the list, which can be expensive if the list is long. An alternative representation is used here in terms of a recursive data structure that can share subparts.

A path is either:

- a node (representing a path of length 0) or
- a path, `initial` and an arc, where the `from_node` of the arc is the node at the end of `initial`.

These cases are distinguished in the following code by having `arc = None` if the path has length 0, in which case `initial` is the node of the path. Python `yield` is used for enumerations only.

```python
class Path(object):
    """A path is either a node or a path followed by an arc""

def __init__(self,initial,arc=None):
    """""initial is either a node (in which case arc is None) or a path (in which case arc is an object of type Arc)""
    self.initial = initial
```

http://aipython.org
3.1. Representing Search Problems

```python
self.arc = arc
if arc is None:
    self.cost = 0
else:
    self.cost = initial.cost + arc.cost

def end(self):
    """returns the node at the end of the path"""
    if self.arc is None:
        return self.initial
    else:
        return self.arc.to_node

def nodes(self):
    """enumerates the nodes for the path.
    This starts at the end and enumerates nodes in the path backwards.""
    current = self
    while current.arc is not None:
        yield current.arc.to_node
        current = current.initial
        yield current.initial

def initial_nodes(self):
    """enumerates the nodes for the path before the end node.
    This starts at the end and enumerates nodes in the path backwards.""
    if self.arc is not None:
        yield from self.initial.nodes()

def __repr__(self):
    """returns a string representation of a path""
    if self.arc is None:
        return str(self.initial)
    elif self.arc.action:
        return "{0} -- {1} \n --> {2}".format(str(self.initial), str(self.arc.action), str(self.arc.to_node))
    else:
        return "{0} --> {1}".format(str(self.initial), str(self.arc.to_node))
```

3.1.3 Example Search Problems

The first search problem is one with 5 nodes where the least-cost path is one with many arcs. See Figure 3.1. Note that this example is used for the unit tests, so the test (in searchGeneric) will need to be changed if this is changed.

```python
problem1 = Search_problem_from_explicit_graph(  
    {'a', 'b', 'c', 'd', 'g'},  
    [Arc('a', 'c', 1), Arc('a', 'b', 3), Arc('c', 'd', 3), Arc('c', 'b', 1),  
```
3. Searching for Solutions

![Figure 3.1: problem1](image)

The second search problem is one with 8 nodes where many paths do not lead to the goal. See Figure 3.2.

```
Arc('b', 'd', 1), Arc('b', 'g', 3), Arc('d', 'g', 1),
start = 'a',
goals = {'g'},
positions={'a': (0, 0), 'b': (1, 1), 'c': (0, 1), 'd': (1, 2), 'g': (2, 2))
```

The third search problem is a disconnected graph (contains no arcs), where the start node is a goal node. This is a boundary case to make sure that weird cases work.

```
problem2 = Search_problem_from_explicit_graph(
  {'a', 'b', 'c', 'd', 'e', 'g', 'h', 'j'},
  [Arc('a', 'b', 1), Arc('b', 'c', 3), Arc('b', 'd', 1), Arc('d', 'e', 3),
    Arc('d', 'g', 1), Arc('a', 'h', 3), Arc('h', 'j', 1)],
  start = 'a',
goals = {'g'},
  positions={'a': (0, 0), 'b': (0, 1), 'c': (0, 4), 'd': (1, 1), 'e': (1, 4),
            'g': (2, 1), 'h': (3, 0), 'j': (3, 1)}
```

http://aipython.org

3.1. Representing Search Problems

The acyclic delivery problem is the delivery problem described in Example 3.4 and shown in Figure 3.2 of the textbook.
The cyclic_delivery_problem is the delivery problem described in Example 3.8 and shown in Figure 3.6 of the textbook. This is the same as cyclic_delivery_problem, but almost every arc also has its inverse.

cyclic_delivery_problem = Search_problem_from_explicit_graph(
    {'mail', 'ts', 'o103', 'o109', 'o111', 'b1', 'b2', 'b3', 'b4', 'c1', 'c2', 'c3',
    'o125', 'o123', 'o119', 'r123', 'storage'},
    [ Arc('ts', 'mail', 6), Arc('mail', 'ts', 6),
      Arc('o103', 'ts', 8), Arc('ts', 'o103', 8),
      Arc('o103', 'b3', 4),
      Arc('o103', 'o109', 12), Arc('o109', 'o103', 12),
      Arc('o109', 'o119', 16), Arc('o119', 'o109', 16),
      Arc('o109', 'o111', 4), Arc('o111', 'o109', 4),
      Arc('b1', 'c2', 3),
      Arc('b1', 'b2', 6), Arc('b2', 'b1', 6),
      Arc('b2', 'b4', 3), Arc('b4', 'b2', 3),
      Arc('b3', 'b1', 4), Arc('b1', 'b3', 4),
      Arc('b3', 'b4', 7), Arc('b4', 'b3', 7),
      Arc('b4', 'o109', 7),
      Arc('c1', 'c3', 8), Arc('c3', 'c1', 8),
      Arc('c2', 'c3', 6), Arc('c3', 'c2', 6),
      Arc('c2', 'c1', 4), Arc('c1', 'c2', 4),
      Arc('o123', 'o103', 4), Arc('o103', 'o123', 4),
      Arc('o123', 'r123', 4), Arc('r123', 'o123', 4),
      Arc('o119', 'o123', 9), Arc('o123', 'o119', 9),
      Arc('o119', 'storage', 7), Arc('storage', 'o119', 7)],
    start = 'o103',
    goals = {'r123'},
    hmap = {
        'mail' : 26,
        'ts' : 23,
        'o103' : 21,
        'o109' : 24,
        'o111' : 27,
        'o119' : 11,
        'o123' : 4,
        'o125' : 6,
        'r123' : 0,
        'b1' : 13,
        'b2' : 15,
        'b3' : 17,
        'b4' : 18,
        'c1' : 6,
        'c2' : 10,
        'c3' : 12,
    })
3.2 Generic Searcher and Variants

To run the search demos, in folder “aipython”, load “searchGeneric.py”, using e.g., `ipython -i searchGeneric.py`, and copy and paste the example queries at the bottom of that file. This requires Python 3.

3.2.1 Searcher

A Searcher for a problem can be asked repeatedly for the next path. To solve a problem, we can construct a Searcher object for the problem and then repeatedly ask for the next path using `search`. If there are no more paths, `None` is returned.

```
from display import Displayable, visualize

class Searcher(Displayable):
    """returns a searcher for a problem.
    Paths can be found by repeatedly calling search().
    This does depth-first search unless overridden """
    def __init__(self, problem):
        """creates a searcher from a problem """
        self.problem = problem
        self.initialize_frontier()
        self.num_expanded = 0
        self.add_to_frontier(Path(problem.start_node()))
        super().__init__()

    def initialize_frontier(self):
        self.frontier = []

    def empty_frontier(self):
        return self.frontier == []

    def add_to_frontier(self, path):
        self.frontier.append(path)

    @visualize
    def search(self):
        """returns (next) path from the problem's start node to a goal node."
```

http://aipython.org  Version 0.9.4  March 2, 2022
Returns None if no path exists.

```python
while not self.empty_frontier():
    path = self.frontier.pop()
    self.display(2, "Expanding:",path,"(cost:",path.cost,")")
    self.num_expanded += 1
    if self.problem.is_goal(path.end()): # solution found
        self.display(1, self.num_expanded, "paths have been expanded
and",
        len(self.frontier), "paths remain in the
        frontier")
        self.solution = path # store the solution found
        return path
    else:
        neighs = self.problem.neighbors(path.end())
        self.display(3,"Neighbors are", neighs)
        for arc in reversed(list(neighs)):
            self.add_to_frontier(Path(path,arc))
        self.display(3,"Frontier:",self.frontier)
        self.display(1,"No (more) solutions. Total of",
        self.num_expanded,"paths expanded.")
```

Note that this reverses the neighbours so that it implements depth-first search in an intuitive manner (expanding the first neighbor first), and list is needed if the neighbours are generated. Reversing the neighbours might not be required for other methods. The calls to reversed and list can be removed, and the algorithm still implements depth-first search.

**Exercise 3.1** When it returns a path, the algorithm can be used to find another path by calling `search()` again. However, it does not find other paths that go through one goal node to another. Explain why, and change the code so that it can find such paths when `search()` is called again.

### 3.2.2 Frontier as a Priority Queue

In many of the search algorithms, such as A* and other best-first searchers, the frontier is implemented as a priority queue. Here we use the Python’s built-in priority queue implementations, `heapq`.

Following the lead of the Python documentation, a frontier is a list of triples. The first element of each triple is the value to be minimized. The second element is a unique index which specifies the order when the first elements are the same, and the third element is the path that is on the queue. The use of the unique index ensures that the priority queue implementation does not compare paths; whether one path is less than another is not defined. It also lets us control what sort of search (e.g., depth-first or breadth-first) occurs when the value to be minimized does not give a unique next path.

[http://aipython.org](http://aipython.org)
The variable `frontier_index` is the total number of elements of the frontier that have been created. As well as being used as a unique index, it is useful for statistics, particularly in conjunction with the current size of the frontier.

```python
import heapq  # part of the Python standard library
from searchProblem import Path

class FrontierPQ(object):
    """A frontier consists of a priority queue (heap), frontierpq, of
    (value, index, path) triples, where
    * value is the value we want to minimize (e.g., path cost + h).
    * index is a unique index for each element
    * path is the path on the queue
    Note that the priority queue always returns the smallest element.
    """

    def __init__(self):
        """constructs the frontier, initially an empty priority queue
        """
        self.frontier_index = 0  # the number of items ever added to the
        frontier
        self.frontierpq = []  # the frontier priority queue

    def empty(self):
        """is True if the priority queue is empty"
        return self.frontierpq == []

    def add(self, path, value):
        """add a path to the priority queue
        value is the value to be minimized"
        self.frontier_index += 1  # get a new unique index
        heapq.heappush(self.frontierpq, (value, -self.frontier_index, path))

    def pop(self):
        """returns and removes the path of the frontier with minimum value.
        """
        (_,_,path) = heapq.heappop(self.frontierpq)
        return path

    def count(self, val):
        """returns the number of elements of the frontier with value=val"
        return sum(1 for e in self.frontierpq if e[0]==val)

    def __repr__(self):
        """string representation of the frontier"
        return str([(n,c,str(p)) for (n,c,p) in self.frontierpq])
```

The following methods are used for finding and printing information about the frontier.
3. Searching for Solutions

```python
def __len__(self):
    """length of the frontier""
    return len(self.frontierpq)

def __iter__(self):
    """iterate through the paths in the frontier""
    for (_, _, path) in self.frontierpq:
        yield path
```

3.2.3 A* Search

For an A* Search the frontier is implemented using the FrontierPQ class.

```python
class AStarSearcher(Searcher):
    """returns a searcher for a problem.
    Paths can be found by repeatedly calling search().
    """
    def __init__(self, problem):
        super().__init__(problem)

    def initialize_frontier(self):
        self.frontier = FrontierPQ()

    def empty_frontier(self):
        return self.frontier.empty()

    def add_to_frontier(self, path):
        """add path to the frontier with the appropriate cost"
        value = path.cost + self.problem.heuristic(path.end())
        self.frontier.add(path, value)
```

Code should always be tested. The following provides a simple unit test, using problem1 as the the default problem.

```python
import searchProblem as searchProblem

def test(SearchClass, problem=searchProblem.problem1,
    solutions=[['g', 'd', 'b', 'c', 'a']]):
    """Unit test for aiPython searching algorithms.
    SearchClass is a class that takes a problem and implements search()
    problem is a search problem
    solutions is a list of optimal solutions
    """
    print("Testing problem 1:")
    schr1 = SearchClass(problem)
    path1 = schr1.search()
```

http://aiPython.org  Version 0.9.4  March 2, 2022
3.2. Generic Searcher and Variants

```python
print("Path found:", path1)
assert path1 is not None, "No path is found in problem1"
assert list(path1.nodes()) in solutions, "Shortest path not found in problem1"
print("Passed unit test")

if __name__ == "__main__":
    #test(Searcher)
    test(AStarSearcher)

    # example queries:
    # searcher1 = Searcher(searchProblem.acyclic_delivery_problem) # DFS
    # searcher1.search() # find first path
    # searcher1.search() # find next path
    # searcher2 = AStarSearcher(searchProblem.acyclic_delivery_problem) # A*
    # searcher2.search() # find first path
    # searcher2.search() # find next path
    # searcher3 = Searcher(searchProblem.cyclic_delivery_problem) # DFS
    # searcher3.search() # find first path with DFS. What do you expect to happen?
    # searcher4 = AStarSearcher(searchProblem.cyclic_delivery_problem) # A*
    # searcher4.search() # find first path
```

**Exercise 3.2** Change the code so that it implements (i) best-first search and (ii) lowest-cost-first search. For each of these methods compare it to $A^*$ in terms of the number of paths expanded, and the path found.

**Exercise 3.3** In the `add` method in `FrontierPQ` what does the "-" in front of `frontier_index` do? When there are multiple paths with the same $f$-value, which search method does this act like? What happens if the "-" is removed? When there are multiple paths with the same value, which search method does this act like? Does it work better with or without the "-"? What evidence did you base your conclusion on?

**Exercise 3.4** The searcher acts like a Python iterator, in that it returns one value (here a path) and then returns other values (paths) on demand, but does not implement the iterator interface. Change the code so it implements the iterator interface. What does this enable us to do?

### 3.2.4 Multiple Path Pruning

The following implements $A^*$ with multiple-path pruning. It overrides `search()` in `Searcher`.

```python
from searchGeneric import AStarSearcher, visualize
from searchProblem import Path
```

To run the multiple-path pruning demo, in folder "aipython", load "searchMPP.py", using e.g., `ipython -i searchMPP.py`, and copy and paste the example queries at the bottom of that file.
Exercise 3.5  Implement a searcher that implements cycle pruning instead of multiple-path pruning. You need to decide whether to check for cycles when paths are added to the frontier or when they are removed. (Hint: either method can be implemented by only changing one or two lines in SearcherMPP. Hint: there is a cycle if path.end() in path.initial_nodes() ) Compare no pruning, multiple path
pruning and cycle pruning for the cyclic delivery problem. Which works better in terms of number of paths expanded, computational time or space?

3.3 Branch-and-bound Search

To run the demo, in folder “aipython”, load “searchBranchAndBound.py”, and copy and paste the example queries at the bottom of that file.

Depth-first search methods do not need an a priority queue, but can use a list as a stack. In this implementation of branch-and-bound search, we call search to find an optimal solution with cost less than bound. This uses depth-first search to find a path to a goal that extends path with cost less than the bound. Once a path to a goal has been found, that path is remembered as the best path, the bound is reduced, and the search continues.

```python
from searchProblem import Path
from searchGeneric import Searcher
from display import Displayable, visualize

class DF_branch_and_bound(Searcher):
    """returns a branch and bound searcher for a problem.
    An optimal path with cost less than bound can be found by calling
    search()
    """
    def __init__(self, problem, bound=float("inf")):
        """creates a searcher than can be used with search() to find an
        optimal path.
        bound gives the initial bound. By default this is infinite -
        meaning there
        is no initial pruning due to depth bound
        """
        super().__init__(problem)
        self.best_path = None
        self.bound = bound
    @visualize
    def search(self):
        """returns an optimal solution to a problem with cost less than
        bound.
        returns None if there is no solution with cost less than bound."
        self.frontier = [Path(self.problem.start_node())]
        self.num_expanded = 0
        while self.frontier:
            path = self.frontier.pop()
            if path.cost+self.problem.heuristic(path.end()) < self.bound:
                # if path.end() not in path.initial_nodes(): # for cycle
                pruning
```

http://aipython.org
Exercise 3.6  Implement a branch-and-bound search uses recursion. Hint: you don’t need an explicit frontier, but can do a recursive call for the children.

Exercise 3.7  After the branch-and-bound search found a solution, Sam ran search again, and noticed a different count. Sam hypothesized that this count was related to the number of nodes that an A* search would use (either expand or be added to the frontier). Or maybe, Sam thought, the count for a number of nodes when the bound is slightly above the optimal path case is related to how A* would work. Is there relationship between these counts? Are there different things that it could count so they are related? Try to find the most specific statement that is true, and explain why it is true.

To test the hypothesis, Sam wrote the following code, but isn’t sure it is helpful:
from searchGeneric import Searcher, AStarSearcher
from searchBranchAndBound import DF_branch_and_bound
from searchMPP import SearcherMPP

DF_branch_and_bound.max_display_level = 1
Searcher.max_display_level = 1

def run(problem,name):
    print("\n\n******",name)

    print("\nA*:")
    asearcher = AStarSearcher(problem)
    print("Path found:", asearcher.search(), " cost=", asearcher.solution.cost)
    print("there are", asearcher.frontier.count(asearcher.solution.cost),
          "elements remaining on the queue with
          f-value=", asearcher.solution.cost)

    print("\nA* with MPP:"),
    msearcher = SearcherMPP(problem)
    print("Path found:", msearcher.search(), " cost=", msearcher.solution.cost)
    print("there are", msearcher.frontier.count(msearcher.solution.cost),
          "elements remaining on the queue with
          f-value=", msearcher.solution.cost)

    bound = asearcher.solution.cost + 0.01
    print("\nBranch and bound (with too-good initial bound of", bound,")")
    tbb = DF_branch_and_bound(problem, bound) # cheating!!!!
    print("Path found:", tbb.search(), " cost=", tbb.solution.cost)
    print("Rerunning B&B")
    print("Path found:", tbb.search())

    bbound = asearcher.solution.cost * 2 + 10
    print("\nBranch and bound (with not-very-good initial bound of", bbound,")")
    tbb2 = DF_branch_and_bound(problem, bbound) # cheating!!!!
    print("Path found:", tbb2.search(), " cost=", tbb2.solution.cost)
    print("Rerunning B&B")
    print("Path found:", tbb2.search())

    print("\nDepth-first search: (Use \^C if it goes on forever")
    tsearcher = Searcher(problem)
    print("Path found:", tsearcher.search(), " cost=", tsearcher.solution.cost)

import searchProblem
from searchTest import run
if __name__ == "__main__":
    run(searchProblem.problem1, "Problem 1")
    # run(searchProblem.acyclic_delivery_problem, "Acyclic Delivery")
    # run(searchProblem.cyclic_delivery_problem, "Cyclic Delivery")
# also test some graphs with cycles, and some with multiple least-cost paths
Chapter 4

Reasoning with Constraints

4.1 Constraint Satisfaction Problems

4.1.1 Variables

A variable consists of a name, a domain and an optional (x,y) position (for displaying). The domain of a variable is a list or a tuple, as the ordering will matter in the representation of constraints.

```python
import random
import matplotlib.pyplot as plt

class Variable(object):
    '''A random variable.
    name (string) - name of the variable
domain (list) - a list of the values for the variable.
Variables are ordered according to their name.
    '''
    def __init__(self, name, domain, position=None):
        '''Variable
        name a string
domain a list of printable values
position of form (x,y)
        '''
        self.name = name # string
        self.domain = domain # list of values
        self.position = position if position else (random.random(),
                         random.random())
        self.size = len(domain)
```

51
4. Reasoning with Constraints

4.1.2 Constraints

A constraint consists of:

- A tuple (or list) of variables is called the **scope**.

- A **condition** is a Boolean function that takes the same number of arguments as there are variables in the scope. The condition must have a **__name__** property that gives a printable name of the function; built-in functions and functions that are defined using `def` have such a property; for other functions you may need to define this property.

- An optional name

- An optional \((x, y)\) position

```python
class Constraint(object):
    """A Constraint consists of
    * scope: a tuple of variables
    * condition: a Boolean function that can applied to a tuple of values
      for variables in scope
    * string: a string for printing the constraints. All of the strings
      must be unique.
    for the variables
    """
    def __init__(self, scope, condition, string=None, position=None):
        self.scope = scope
        self.condition = condition
        if string is None:
            self.string = self.condition.__name__ + str(self.scope)
        else:
            self.string = string
        self.position = position

    def __repr__(self):
        return self.string
```

An **assignment** is a `variable:value` dictionary.

If `con` is a constraint, `con.holds(assignment)` returns True or False depending on whether the condition is true or false for that assignment. The assignment `assignment` must assigns a value to every variable in the scope of the constraint `con` (and could also assign values other variables); `con.holds` gives an error if
not all variables in the scope of con are assigned in the assignment. It ignores variables in assignment that are not in the scope of the constraint.

In Python, the * notation is used for unpacking a tuple. For example, \( F(1, 2, 3) \) is the same as \( F(1, 2, 3) \). So if \( t \) has value \( (1, 2, 3) \), then \( F(*t) \) is the same as \( F(1, 2, 3) \).

def can_evaluate(self, assignment):
    """
    assignment is a variable:value dictionary
    returns True if the constraint can be evaluated given assignment
    """
    return all(v in assignment for v in self.scope)

def holds(self, assignment):
    """
    returns the value of Constraint con evaluated in assignment.
    precondition: all variables are assigned in assignment, ie
    self.can_evaluate(assignment) is true
    """
    return self.condition(*tuple(assignment[v] for v in self.scope))

4.1.3 CSPs

A constraint satisfaction problem (CSP) requires:

- **variables**: a list or set of variables
- **constraints**: a set or list of constraints.

Other properties are inferred from these:

- **variables** is the set of variables.
- **var_to_const** is a mapping from variables to set of constraints, such that
  \( \text{var_to_const}[\text{var}] \) is the set of constraints with \( \text{var} \) in the scope.

class CSP(object):
    """A CSP consists of
    * a title (a string)
    * variables, a set of variables
    * constraints, a list of constraints
    * var_to_const, a variable to set of constraints dictionary
    """
    def __init__(self, title, variables, constraints):
        """title is a string
        variables is set of variables
        constraints is a list of constraints
        """
Reasoning with Constraints

self.title = title
self.variables = variables
self.constraints = constraints
self.var_to_const = {var: set() for var in self.variables}
for con in self.constraints:
    for var in con.scope:
        self.var_to_const[var].add(con)

def __str__(self):
    return str(self.title)

def __repr__(self):
    return f"CSP({self.title}, {self.variables}, {(str(c) for c in self.constraints)})"

csp.consistent(assignment) returns true if the assignment is consistent with each of the constraints in csp (i.e., all of the constraints that can be evaluated evaluate to true). Note that this is a local consistency with each constraint; it does not imply the CSP is consistent or has a solution.

def consistent(self, assignment):

    returns True if all of the constraints that can be evaluated evaluate to True given assignment.

    return all(con.holds(assignment)
               for con in self.constraints
               if con.can_evaluate(assignment))

The show method uses matplotlib to show the graphical structure of a constraint network.

def show(self):
    ...
4.1. Constraint Satisfaction Problems

```python
for i in range(2):
    bbox = dict(boxstyle="square,pad=1.0",color="green")
    for var in con.scope:
        ax.annotate(con.string, var.position, xytext=con.position,
                    arrowprops={'arrowstyle': '='},bbox=con_bbox,
                    ha='center')
    for var in self.variables:
        x,y = var.position
        plt.text(x,y,var.name,bbox=var_bbox,ha='center')
```

4.1.4 Examples

In the following code `ne_`, when given a number, returns a function that is true when its argument is not that number. For example, if \( f = ne_3 \), then \( f(2) \) is True and \( f(3) \) is False. That is, \( ne_3(x)(y) \) is true when \( x \neq y \). Allowing a function of multiple arguments to use its arguments one at a time is called currying, after the logician Haskell Curry. Functions used as conditions in constraints require names (so they can be printed).

```python
def ne_(val):
    """not equal value""
    # nev = lambda x: x != val # alternative definition
    # nev = partial(neq,val) # another alternative definition
    def nev(x):
        return val != x
    nev.__name__ = str(val)+"=" # name of the function
    return nev
```

Similarly \( is_(x)(y) \) is true when \( x = y \).

```python
def is_(val):
    """is a value""
    # isv = lambda x: x == val # alternative definition
    # isv = partial(eq,val) # another alternative definition
    def isv(x):
        return x == val
    isv.__name__ = str(val)+"="
    return isv
```

The CSP, \( csp_0 \) has variables \( X \), \( Y \) and \( Z \), each with domain \{1, 2, 3\}. The constraints are \( X < Y \) and \( Y < Z \).
The CSP, csp1 has variables A, B and C, each with domain \{1, 2, 3, 4\}. The constraints are \(A < B\), \(B \neq 2\) and \(B < C\). This is slightly more interesting than csp0 as it has more solutions. This example is used in the unit tests, and so if it is changed, the unit tests need to be changed.

The next CSP, csp2 is Example 4.9 of the textbook; the domain consistent network (after applying the unary constraints) is shown in Figure ??  . Note that we use the same variables as the previous example and add two more.
4.1. Constraint Satisfaction Problems

The following example is another scheduling problem (but with multiple answers). This is the same a scheduling 2 in the original Alispaces.org consistency app.

```
csp3 = CSP("csp3", {A,B,C,D,E},
[Constraint([A,B], ne, "A != B"),
 Constraint([A,D], lt, "A < D"),
 Constraint([A,E], lambda a,e: (a-e)%2 == 1, "A-E is odd"),
 Constraint([B,E], lt, "B < E"),
 Constraint([D,C], lt, "D < C"),
 Constraint([C,E], ne, "C != E"),
 Constraint([D,E], ne, "D != E")]
```

http://aipython.org
Version 0.9.4
March 2, 2022
The following example is another abstract scheduling problem. What are the solutions?

```python
def adjacent(x, y):
    """True when x and y are adjacent numbers"""
    return abs(x - y) == 1

csp4 = CSP("csp4", {A, B, C, D, E},
               [Constraint([A, B], adjacent, "adjacent(A,B)")],
               [Constraint([B, C], adjacent, "adjacent(B,C)")],
               [Constraint([C, D], adjacent, "adjacent(C,D)")],
               [Constraint([D, E], adjacent, "adjacent(D,E)")],
               [Constraint([A, C], ne, "A != C"),
                Constraint([B, D], ne, "B != D"),
                Constraint([C, E], ne, "C != E")])
```

The following examples represent the crossword shown in Figure 4.5.

In the first representation, the variables represent words. The constraint imposed by the crossword is that where two words intersect, the letter at the intersection must be the same. The method `meet_at` is used to test whether two words intersect with the same letter. For example, the constraint `meet_at(2, 0)`
4.1. Constraint Satisfaction Problems

Figure 4.4: csp4.show()

Figure 4.5: crossword1: a crossword puzzle to be solved

Words:
ant, big, bus, car, has, book, buys, hold, lane, year, ginger, search, symbol, syntax.
means that the third letter (at position 2) of the first argument is the same as the first letter of the second argument. This is shown in Figure 4.6.

def meet_at(p1, p2):
    r"""returns a function of two words that is true when the words intersect at positions p1, p2.
    The positions are relative to the words; starting at position 0.
    meet_at(p1,p2)(w1,w2) is true if the same letter is at position p1 of word w1
    and at position p2 of word w2.
    ""
    def meets(w1, w2):
        return w1[p1] == w2[p2]
    meets.__name__ = "meet_at(\"+str(p1)+\",\"+str(p2)+\")"
    return meets

one_across = Variable('one_across', {'ant', 'big', 'bus', 'car', 'has'}, position=(0.3,0.9))
one_down = Variable('one_down', {'book', 'buys', 'hold', 'lane', 'year'}, position=(0.1,0.7))
two_down = Variable('two_down', {'ginger', 'search', 'symbol', 'syntax'}, position=(0.9,0.8))
three_across = Variable('three_across', {'book', 'buys', 'hold', 'land', 'year'}, position=(0.1,0.3))
four_across = Variable('four_across', {'ant', 'big', 'bus', 'car', 'has'}, position=(0.7,0.0))
4.1. Constraint Satisfaction Problems

In an alternative representation of a crossword (the “dual” representation), the variables represent letters, and the constraints are that adjacent sequences of letters form words. This is shown in Figure 4.7.

crossword1 = CSP("crossword1",
    {one_across, one_down, two_down, three_across,
        four_across},
    [Constraint([one_across,one_down], meet_at(0,0)),
     Constraint([one_across,two_down], meet_at(2,0)),
     Constraint([three_across,two_down], meet_at(2,2)),
     Constraint([three_across,one_down], meet_at(0,2)),
     Constraint([four_across,two_down], meet_at(0,4))])

def is_word(*letters, words=words):
    """is true if the letters concatenated form a word in words""
    return """.join(letters) in words

letters = {'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'k', 'l',
    'm', 'n', 'o', 'p', 'q', 'r', 's', 't', 'u', 'v', 'w', 'x', 'y',
    'z'}

words = {'ant', 'big', 'bus', 'car', 'has', 'book', 'buys', 'hold',
    'lane', 'year', 'ginger', 'search', 'symbol', 'syntax'}

Figure 4.7: crossword1d.show()
"z"}

# pij is the variable representing the letter i from the left and j down
(starting from 0)
p00 = Variable('p00', letters, position=(0.1,0.85))
p10 = Variable('p10', letters, position=(0.3,0.85))
p20 = Variable('p20', letters, position=(0.5,0.85))
p01 = Variable('p01', letters, position=(0.1,0.7))
p21 = Variable('p21', letters, position=(0.5,0.7))
p02 = Variable('p02', letters, position=(0.1,0.55))
p12 = Variable('p12', letters, position=(0.3,0.55))
p22 = Variable('p22', letters, position=(0.5,0.55))
p32 = Variable('p32', letters, position=(0.7,0.55))
p03 = Variable('p03', letters, position=(0.1,0.4))
p23 = Variable('p23', letters, position=(0.5,0.4))
p24 = Variable('p24', letters, position=(0.5,0.25))
p34 = Variable('p34', letters, position=(0.7,0.25))
p44 = Variable('p44', letters, position=(0.9,0.25))
p25 = Variable('p25', letters, position=(0.5,0.1))

crossword1d = CSP("crossword1d",
    {p00, p10, p20, # first row
     p01, p21, # second row
     p02, p12, p22, p32, # third row
     p03, p23, #fourth row
     p24, p34, p44, # fifth row
     p25 # sixth row
    },
    [Constraint([p00, p10, p20], is_word,
                position=(0.3,0.95)), #1-across
     Constraint([p00, p01, p02, p03], is_word,
                position=(0,0.625)), # 1-down
     Constraint([p02, p12, p22, p32], is_word,
                position=(0.3,0.625)), # 3-across
     Constraint([p20, p21, p22, p23, p24, p25], is_word,
                position=(0.45,0.475)), # 2-down
     Constraint([p24, p34, p44], is_word,
                position=(0.7,0.325)) # 4-across
    ])

Exercise 4.1  How many assignments of a value to each variable are there for
each of the representations of the above crossword? Do you think an exhaustive
enumeration will work for either one?

The queens problem is a puzzle on a chess board, where the idea is to place
a queen on each column so the queens cannot take each other: there are no
two queens on the same row, column or diagonal. The n-queens problem is a
generalization where the size of the board is an n × n, and n queens have to be
placed.

http://aipython.org  Version 0.9.4  March 2, 2022
4.1. Constraint Satisfaction Problems

Here is a representation of the n-queens problem, where the variables are the columns and the values are the rows in which the queen is placed. The original queens problem on a standard \((8 \times 8)\) chess board is \(n\_queens(8)\).

```python
def queens(ri,rj):
    """ri and rj are different rows, return the condition that the queens cannot take each other""
    def no_take(ci,cj):
        """is true if queen at (ri,ci) cannot take a queen at (rj,cj)""
        return ci != cj and abs(ri-ci) != abs(rj-cj)
    return no_take

def n_queens(n):
    """returns a CSP for n-queens""
    columns = list(range(n))
    variables = [Variable(f"R{i}",columns) for i in range(n)]
    return CSP("n-queens",
                variables,
                [Constraint([variables[i], variables[j]], queens(i,j))
                 for i in range(n) for j in range(n) if i != j])

# try the CSP n_queens(8) in one of the solvers.
# What is the smallest n for which there is a solution.
```

**Exercise 4.2** How many constraints does this representation of the n-queens problem produce? Can it be done with fewer constraints? Either explain why it can’t be done with fewer constraints, or give a solution using fewer constraints.

**Unit tests**

The following defines a **unit test** for csp solvers, by default using example csp1.

```python
def test_csp(CSP_solver, csp=csp1,
            solutions=[{A: 1, B: 3, C: 4}, {A: 2, B: 3, C: 4}]):
    """CSP_solver is a solver that takes a csp and returns a solution
    csp is a constraint satisfaction problem
    solutions is the list of all solutions to csp
    This tests whether the solution returned by CSP_solver is a solution.
    """
    print("Testing csp with",CSP_solver.__doc__)
    sol0 = CSP_solver(csp)
    print("Solution found:",sol0)
    assert sol0 in solutions, "Solution not correct for " + str(csp)
    print("Passed unit test")
```

**Exercise 4.3** Modify test so that instead of taking in a list of solutions, it checks whether the returned solution actually is a solution.
Exercise 4.4  Propose a test that is appropriate for CSPs with no solutions. Assume that the test designer knows there are no solutions. Consider what a CSP solver should return if there are no solutions to the CSP.

Exercise 4.5  Write a unit test that checks whether all solutions (e.g., for the search algorithms that can return multiple solutions) are correct, and whether all solutions can be found.

4.2  A Simple Depth-first Solver

The first solver searches through the space of partial assignments. This takes in a CSP problem and an optional variable ordering, which is a list of the variables in the CSP. It returns a generator of the solutions (see Python documentation on yield for enumerations).

cspDFS.py — Solving a CSP using depth-first search.

```python
from cspExamples import csp1, csp2, test_csp, crossword1, crossword1d

def dfs_solver(constraints, context, var_order):
    """generator for all solutions to csp.
    context is an assignment of values to some of the variables.
    var_order is a list of the variables in csp that are not in context.
    """
    to_eval = {c for c in constraints if c.can_evaluate(context)}
    if all(c.holds(context) for c in to_eval):
        if var_order == []:
            yield context
        else:
            rem_cons = [c for c in constraints if c not in to_eval]
            var = var_order[0]
            for val in var.domain:
                yield from dfs_solver(rem_cons, context | {var: val}, var_order[1:1])

def dfs_solve_all(csp, var_order=None):
    """depth-first CSP solver to return a list of all solutions to csp.
    """
    if var_order == None:  # use an arbitrary variable order
        var_order = list(csp.variables)
    return list(dfs_solver(csp.constraints, {}, var_order))

def dfs_solve1(csp, var_order=None):
    """depth-first CSP solver to find single solution or None if there are
    no solutions.
    """
    if var_order == None:  # use an arbitrary variable order
        var_order = list(csp.variables)
    gen = dfs_solver(csp.constraints, {}, var_order)
    try:  # Python generators raise an exception if there are no more elements.
        return next(gen)
    except StopIteration:
        return None
```

http://aipython.org
4.3. Converting CSPs to Search Problems

4.3.1 return next(gen)
        except StopIteration:
            return None

        if __name__ == "__main__":
            test_csp(dfs_solve1)

        # Try:
        # dfs_solve_all(csp1)
        # dfs_solve_all(csp2)
        # dfs_solve_all(crossword1)
        # dfs_solve_all(crossword1d) # warning: may take a *very* long time!

Exercise 4.6 Instead of testing all constraints at every node, change it so each constraint is only tested when all of it variables are assigned. Given an elimination ordering, it is possible to determine when each constraint needs to be tested. Implement this. Hint: create a parallel list of sets of constraints, where at each position $i$ in the list, the constraints at position $i$ can be evaluated when the variable at position $i$ has been assigned.

Exercise 4.7 Estimate how long dfs_solve_all(crossword1d) will take on your computer. To do this, reduce the number of variables that need to be assigned, so that the simplifies problem can be solved in a reasonable time (between 0.1 second and 10 seconds). This can be done by reducing the number of variables in var_order, as the program only splits on these. How much more time will it take if the number of variables is increased by 1? (Try it!) Then extrapolate to all of the variables. See Section 1.6.1 for how to time your code. Would making the code 100 times faster or using a computer 100 times faster help?

4.3 Converting CSPs to Search Problems

To run the demo, in folder "aipython", load "cspSearch.py", and copy and paste the example queries at the bottom of that file.

The next solver constructs a search space that can be solved using the search methods of the previous chapter. This takes in a CSP problem and an optional variable ordering, which is a list of the variables in the CSP. In this search space:

- A node is a variable : value dictionary which does not violate any constraints (so that dictionaries that violate any constraints are not added).

- An arc corresponds to an assignment of a value to the next variable. This assumes a static ordering; the next variable chosen to split does not depend on the context. If no variable ordering is given, this makes no attempt to choose a good ordering.

http://aipython.org Version 0.9.4 March 2, 2022
4. Reasoning with Constraints

from cspProblem import CSP, Constraint
from searchProblem import Arc, Search_problem
from utilities import dict_union

class Search_from_CSP(Search_problem):
    """A search problem directly from the CSP."
    
    A node is a variable:value dictionary"
    def __init__(self, csp, variable_order=None):
        self.csp=csp
        if variable_order:
            assert set(variable_order) == set(csp.variables)
            assert len(variable_order) == len(csp.variables)
            self.variables = variable_order
        else:
            self.variables = list(csp.variables)

    def is_goal(self, node):
        """returns whether the current node is a goal for the search"
        return len(node)==len(self.csp.variables)

    def start_node(self):
        """returns the start node for the search"
        return {}

    def neighbors(self, node):
        """returns a list of the neighboring nodes of node."
        var = self.variables[len(node)] # the next variable
        res = []
        for val in var.domain:
            new_env = dict_union(node,{var:val}) #dictionary union
            if self.csp.consistent(new_env):
                res.append(Arc(node,new_env))
        return res

The `neighbors(node)` method uses the fact that the length of the node, which is the number of variables already assigned, is the index of the next variable to split on. Note that we do no need to check whether there are no more variables to split on, as the nodes are all consistent, by construction, and so when there are no more variables we have a solution, and so don’t need the neighbours.

The unit tests relies on a solver. The following procedure creates a solver using search that can be tested.
from searchGeneric import Searcher

def solver_from_searcher(csp):
    
    """depth-first search solver""
    path = Searcher(Search_from_CSP(csp)).search()
    if path is not None:
        return path.end()
    else:
        return None

if __name__ == "__main__":
    test_csp(solver_from_searcher)

## Test Solving CSPs with Search:
searcher1 = Searcher(Search_from_CSP(csp1))
# print(searcher1.search()) # get next solution
searcher2 = Searcher(Search_from_CSP(csp2))
# print(searcher2.search()) # get next solution
searcher3 = Searcher(Search_from_CSP(crossword1))
# print(searcher3.search()) # get next solution
searcher4 = Searcher(Search_from_CSP(crossword1d))
# print(searcher4.search()) # get next solution (warning: slow)

Exercise 4.8 What would happen if we constructed the new assignment by assigning node[var] = val (with side effects) instead of using dictionary union? Give an example of where this could give a wrong answer. How could the algorithm be changed to work with side effects? (Hint: think about what information needs to be in a node).

Exercise 4.9 Change neighbors so that it returns an iterator of values rather than a list. (Hint: use yield.)

4.4 Consistency Algorithms

A Con_solver is used to simplify a CSP using arc consistency.

To run the demo, in folder "aipython", load "cspConsistency.py", and copy and paste the commented-out example queries at the bottom of that file.

http://aipython.org
* kwargs is the keyword arguments for Displayable superclass

```python
self.csp = csp
super().__init__(**kwargs) # Or Displayable.__init__(self,**kwargs)
```

The following implementation of arc consistency maintains the set to_do of (variable, constraint) pairs that are to be checked. It takes in a domain dictionary and returns a new domain dictionary. It needs to be careful to avoid side effects (by copying the domains dictionary and the to_do set).

```python
def make_arc_consistent(self, orig_domains=None, to_do=None):
    """Makes this CSP arc-consistent using generalized arc consistency
    orig_domains is the original domains
    to_do is a set of (variable,constraint) pairs
    returns the reduced domains (an arc-consistent variable:domain
    dictionary)
    """
    if orig_domains is None:
        orig_domains = {var:var.domain for var in self.csp.variables}
    if to_do is None:
        to_do = {(var, const) for const in self.csp.constraints
        for var in const.scope}
    else:
        to_do = to_do.copy() # use a copy of to_do
        domains = orig_domains.copy()
        self.display(2,"Performing AC with domains", domains)
        while to_do:
            var, const = self.select_arc(to_do)
            self.display(3, "Processing arc (", var, ",", const, ")")
            other_vars = [ov for ov in const.scope if ov != var]
            new_domain = {val for val in domains[var]
                          if self.any_holds(domains, const, {var: val},
                          other_vars)}
            if new_domain != domains[var]:
                self.display(4, "Arc: (", var, ",", const, ") is
                inconsistent")
                self.display(3, "Domain pruned", "dom("var, ") =",
                new_domain,
                " due to ", const)
                domains[var] = new_domain
                add_to_do = self.new_to_do(var, const) - to_do
                to_do |= add_to_do # set union
                self.display(3, " adding", add_to_do if add_to_do else
                "nothing", "to to.do.")
                self.display(4, "Arc: (", var, ",", const, ") now consistent")
        self.display(2, "AC done. Reduced domains", domains)
        return domains

def new_to_do(self, var, const):
    """returns new elements to be added to to_do after assigning
```
variable var in constraint const.

```python
return {(nvar, nconst) for nconst in self.csp.var_to_const[var]
        if nconst != const
        for nvar in nconst.scope
        if nvar != var}
```

The following selects an arc. Any element of `to_do` can be selected. The selected element needs to be removed from `to_do`. The default implementation just selects which ever element `pop` method for sets returns. A user interface could allow the user to select an arc. Alternatively a more sophisticated selection could be employed (or just a stack or a queue).

```python
def select_arc(self, to_do):
    """Selects the arc to be taken from to_do.
    * to_do is a set of arcs, where an arc is a (variable,constraint) pair
    the element selected must be removed from to_do.
    """
    return to_do.pop()
```

The value of `new_domain` is the subset of the domain of `var` that is consistent with the assignment to the other variables. It might be easier to understand the following code, which treats unary (with no other variables in the constraint) and binary (with one other variables in the constraint) constraints as special cases (this can replace the assignment to `new_domain` in the above code):

```python
if len(other_vars)==0:  # unary constraint
    new_domain = {val for val in domains[var]
                  if const.holds({var:val})}
elif len(other_vars)==1:  # binary constraint
    other = other_vars[0]
    new_domain = {val for val in domains[var]
                  if any(const.holds({var: val,other:other_val})
                       for other_val in domains[other])}
else:  # general case
    new_domain = {val for val in domains[var]
                  if self.any_holds(domains, const, {var: val}, other_vars)}
```

`any_holds` is a recursive function that tries to finds an assignment of values to the other variables (`other_vars`) that satisfies constraint `const` given the assignment in `env`. The integer variable `ind` specifies which index to `other_vars` needs to be checked next. As soon as one assignment returns `True`, the algorithm returns `True`. Note that it has side effects with respect to `env`; it changes the values of the variables in `other_vars`. It should only be called when the side effects have no ill effects.
4. Reasoning with Constraints

```python
def any_holds(self, domains, const, env, other_vars, ind=0):
    """returns True if Constraint const holds for an assignment
    that extends env with the variables in other_vars[ind:]
    env is a dictionary
    Warning: this has side effects and changes the elements of env
    """
    if ind == len(other_vars):
        return const.holds(env)
    else:
        var = other_vars[ind]
        for val in domains[var]:
            env[var] = val
            if self.any_holds(domains, const, env, other_vars, ind + 1):
                return True
    return False
```

4.4.1 Direct Implementation of Domain Splitting

The following is a direct implementation of domain splitting with arc consistency that uses recursion. It finds one solution if one exists or returns False if there are no solutions.

```python
def solve_one(self, domains=None, to_do=None):
    """return a solution to the current CSP or False if there are no
    solutions
to_do is the list of arcs to check
    """
    new_domains = self.make_arc_consistent(domains, to_do)
    if any(len(new_domains[var]) == 0 for var in new_domains):
        return False
    elif all(len(new_domains[var]) == 1 for var in new_domains):
        self.display(2, "solution:", {var: select(new_domains[var]) for var in new_domains})
        return {var: select(new_domains[var]) for var in new_domains}
    else:
        var = self.select_var(x for x in self.csp.variables if len(new_domains[x]) > 1)
        if var:
            dom1, dom2 = partition_domain(new_domains[var])
            self.display(3, "...splitting", var, "into", dom1, "and", dom2)
            new_doms1 = copy_with_assign(new_domains, var, dom1)
            new_doms2 = copy_with_assign(new_domains, var, dom2)
            to_do = self.new_to_do(var, None)
            self.display(3, " adding", to_do if to_do else "nothing", "to to_do.")
            return self.solve_one(new_doms1, to_do) or self.solve_one(new_doms2, to_do)
```

http://aipython.org  Version 0.9.4  March 2, 2022
4.4. Consistency Algorithms

```python
def select_var(self, iter_vars):
    """return the next variable to split""
    return select(iter_vars)

def partition_domain(dom):
    """partitions domain dom into two."
    split = len(dom) // 2
    dom1 = set(list(dom)[:split])
    dom2 = dom - dom1
    return dom1, dom2
```

The domains are implemented as a dictionary that maps each variables to its domain. Assigning a value in Python has side effects which we want to avoid. `copy_with_assign` takes a copy of the domains dictionary, perhaps allowing for a new domain for a variable. It creates a copy of the CSP with an (optional) assignment of a new domain to a variable. Only the domains are copied.

```python
def copy_with_assign(domains, var=None, new_domain={True, False}):
    """create a copy of the domains with an assignment var=new_domain
    if var==None then it is just a copy."
    newdoms = domains.copy()
    if var is not None:
        newdoms[var] = new_domain
    return newdoms
```

```python
def select(iterable):
    """select an element of iterable. Returns None if there is no such element.
    This implementation just picks the first element.
    For many of the uses, which element is selected does not affect correctness,
    but may affect efficiency."
    for e in iterable:
        return e # returns first element found
```

**Exercise 4.10** Implement of `solve_all` that is like `solve_one` but returns the set of all solutions.

**Exercise 4.11** Implement `solve_enum` that enumerates the solutions. It should use Python’s `yield` (and perhaps `yield from`).

Unit test:

*http://aipython.org*
4. Reasoning with Constraints

4.4.2 Domain Splitting as an interface to graph searching

An alternative implementation is to implement domain splitting in terms of the search abstraction of Chapter[3].

A node is a domains dictionary.
Exercise 4.12  When splitting a domain, this code splits the domain into half, approximately in half (without any effort to make a sensible choice). Does it work better to split one element from a domain?

Unit test:

```python
from cspExamples import test_csp
from searchGeneric import Searcher

def ac_search_solver(csp):
    """arc consistency (search interface)""
    sol = Searcher(Search_with_AC_from_CSP(csp)).search()
    if sol:
        return {v:select(d) for (v,d) in sol.end().items()}
    if __name__ == '__main__':
        test_csp(ac_search_solver)

Testing:

```
Given a CSP, the stochastic local searcher (SLSearcher) creates the data structures:

- **variables_to_select** is the set of all of the variables with domain-size greater than one. For a variable not in this set, we cannot pick another value from that variable.

- **var_to_constraints** maps from a variable into the set of constraints it is involved in. Note that the inverse mapping from constraints into variables is part of the definition of a constraint.

```python
cspSLS.py — Stochastic Local Search for Solving CSPs

from cspProblem import CSP, Constraint
from searchProblem import Arc, Search_problem
from display import Displayable
import random
import heapq

class SLSearcher(Displayable):
    """A search problem directly from the CSP."
    def __init__(self, csp):
        self.csp = csp
        self.variables_to_select = {var for var in self.csp.variables if len(var.domain) > 1}
        self.current_assignment = None  # this will trigger a random restart
        self.number_of_steps = 0  # number of steps after the initialization

    def restart(self):
        """creates a new total assignment and the conflict set"
        self.current_assignment = {var: random_choice(var.domain) for var in self.csp.variables}
        self.display(2, "Initial assignment", self.current_assignment)
        self.conflicts = set()  # Create assignment and conflicts set
        self.number_of_steps = 0  # number of steps after the initialization

        for con in self.csp.constraints:
            if not con.holds(self.current_assignment):
                self.conflicts.add(con)
        self.display(2, "Number of conflicts", len(self.conflicts))
        self.number_of_steps = 0
```

**restart** creates a new total assignment, and constructs the set of conflicts (the constraints that are false in this assignment).

The **search** method is the top-level searching algorithm. It can either be used to start the search or to continue searching. If there is no current assignment,
it must create one. Note that, when counting steps, a restart is counted as one step.

This method selects one of two implementations. The argument prob_best is the probability of selecting a best variable (one involving the most conflicts). When the value of prob_best is positive, the algorithm needs to maintain a priority queue of variables and the number of conflicts (using search_with_var_pq). If the probability of selecting a best variable is zero, it does not need to maintain this priority queue (as implemented in search_with_any_conflict).

The argument prob_anycon is the probability that the any-conflict strategy is used (which selects a variable at random that is in a conflict), assuming that it is not picking a best variable. Note that for the probability parameters, any value less than zero acts like probability zero and any value greater than 1 acts like probability 1. This means that when prob_anycon = 1.0, a best variable is chosen with probability prob_best, otherwise a variable in any conflict is chosen. A variable is chosen at random with probability 1 − prob_anycon − prob_best as long as that is positive.

This returns the number of steps needed to find a solution, or None if no solution is found. If there is a solution, it is in self.current_assignment.

```python
def search(self, max_steps, prob_best=0, prob_anycon=1.0):
    """
    returns the number of steps or None if there is no solution.
    If there is a solution, it can be found in self.current_assignment
    max_steps is the maximum number of steps it will try before giving up
    prob_best is the probability that a best variable (one in most conflict) is selected
    prob_anycon is the probability that a variable in any conflict is selected
    (otherwise a variable is chosen at random)
    """
    if self.current_assignment is None:
        self.restart()
        self.number_of_steps += 1
    if not self.conflicts:
        self.display(1, "Solution found:", self.current_assignment,
                     "after restart")
        return self.number_of_steps
    if prob_best > 0: # we need to maintain a variable priority queue
        return self.search_with_var_pq(max_steps, prob_best,
                                        prob_anycon)
    else:
        return self.search_with_any_conflict(max_steps, prob_anycon)
```

**Exercise 4.13** This does an initial random assignment but does not do any random restarts. Implement a searcher that takes in the maximum number of walk
steps (corresponding to existing max_steps) and the maximum number of restarts, and returns the total number of steps for the first solution found. (As in search, the solution found can be extracted from the variable self.current_assignment).

### 4.5.1 Any-conflict

If the probability of picking a best variable is zero, the implementation need to keeps track of which variables are in conflicts.

```python
def search_with_any_conflict(self, max_steps, prob_anycon=1.0):
    """Searches with the any_conflict heuristic.
    This relies on just maintaining the set of conflicts;
    it does not maintain a priority queue
    ""
    self.variable_pq = None # we are not maintaining the priority queue.
    # This ensures it is regenerated if
    # we call search_with_var_pq.

    for i in range(max_steps):
        self.number_of_steps +=1
        if random.random() < prob_anycon:
            con = random_choice(self.conflicts) # pick random conflict
            var = random_choice(con.scope) # pick variable in conflict
        else:
            var = random_choice(self.variables_to_select)
        if len(var.domain) > 1:
            val = random_choice([val for val in var.domain
                                 if val is not self.current_assignment[var]])
            self.display(2,self.number_of_steps,"Assigning",var,"=",val)
            self.current_assignment[var]=val
            for varcon in self.csp.var_to_const[var]:
                if varcon.holds(self.current_assignment):
                    self.conflicts.remove(varcon)
                    if varcon not in self.conflicts:
                        self.conflicts.add(varcon)
                        self.display(2," Number of conflicts",len(self.conflicts))
                        if not self.conflicts:
                            self.display(1,"Solution found:",self.current_assignment,
                                         "in", self.number_of_steps,"steps")
                            return self.number_of_steps
                        self.display(1,"No solution in",self.number_of_steps,"steps",
                                     len(self.conflicts),"conflicts remain")
                else:
                    self.conflicts.remove(varcon)
                    self.display(2," Number of conflicts",len(self.conflicts))
                    if not self.conflicts:
                        self.display(1,"Solution found:",self.current_assignment,
                                     "in", self.number_of_steps,"steps")
                        return self.number_of_steps
                    self.display(1,"No solution in",self.number_of_steps,"steps",
                                 len(self.conflicts),"conflicts remain")
        return None
```

**Exercise 4.14** This makes no attempt to find the best alternative value for a variable. Modify the code so that after selecting a variable it selects a value the reduces
4.5. Solving CSPs using Stochastic Local Search

the number of conflicts by the most. Have a parameter that specifies the probability that the best value is chosen.

4.5.2 Two-Stage Choice

This is the top-level searching algorithm that maintains a priority queue of variables ordered by (the negative of) the number of conflicts, so that the variable with the most conflicts is selected first. If there is no current priority queue of variables, one is created.

The main complexity here is to maintain the priority queue. This uses the dictionary var.differential which specifies how much the values of variables should change. This is used with the updatable queue (page 79) to find a variable with the most conflicts.

cspSLS.py — (continued)

```python
def search_with_var_pq(self,max_steps, prob_best=1.0, prob_anycon=1.0):
    """search with a priority queue of variables.
    This is used to select a variable with the most conflicts.
    ""
    if not self.variable_pq:
        self.create_pq()
    pick_best_or_con = prob_best + prob_anycon
    for i in range(max_steps):
        randnum = random.random()
        ## Pick a variable
        if randnum < prob_best: # pick best variable
            var,oldval = self.variable_pq.top()
        elif randnum < pick_best_or_con: # pick a variable in a conflict
            con = random_choice(self.conflicts)
            var = random_choice(con.scope)
        else: #pick any variable that can be selected
            var = random_choice(self.variables_to_select)
        if len(var.domain) > 1: # var has other values
            ## Pick a value
            val = random_choice([val for val in var.domain if val is not
                                  self.current_assignment[var]])
            self.display(2,"Assigning",var,val)
            ## Update the priority queue
            var_differential = {}
            self.current_assignment[var]=val
            for varcon in self.csp.var_to_const[var]:
                self.display(3,"Checking",varcon)
                if varcon.holds(self.current_assignment):
                    if varcon in self.conflicts: # was incons, now consis
                        self.display(3,"Became consistent",varcon)
                        self.conflicts.remove(varcon)
                    for v in varcon.scope: # v is in one fewer
                        conflicts
```

http://aipython.org
var_differential[v] =
var_differential.get(v,0)-1
else:
    if varcon not in self.conflicts: # was consis, not now
        self.display(3,"Became inconsistent",varcon)
        self.conflicts.add(varcon)
        for v in varcon.scope: # v is in one more
            var_differential[v] =
                var_differential.get(v,0)+1
        self.variable_pq.update_each_priority(var_differential)
    if not self.conflicts: # no conflicts, so solution found
        self.display(1,"Solution found:",
            self.current_assignment,"in",
            self.number_of_steps,"steps")
        return self.number_of_steps
        self.display(1,"No solution in",self.number_of_steps,"steps",
            len(self.conflicts),"conflicts remain")
    return None

create_pq creates an updatable priority queue of the variables, ordered by the
number of conflicts they participate in. The priority queue only includes vari-
ables in conflicts and the value of a variable is the negative of the number
of conflicts the variable is in. This ensures that the priority queue, which picks
the minimum value, picks a variable with the most conflicts.
4.5. Solving CSPs using Stochastic Local Search

Exercise 4.15 This makes no attempt to find the best alternative value for a variable. Modify the code so that after selecting a variable it selects a value that reduces the number of conflicts by the most. Have a parameter that specifies the probability that the best value is chosen.

Exercise 4.16 These implementations always select a value for the variable selected that is different from its current value (if that is possible). Change the code so that it does not have this restriction (so it can leave the value the same). Would you expect this code to be faster? Does it work worse (or better)?

4.5.3 Updatable Priority Queues

An updatable priority queue is a priority queue, where key-value pairs can be stored, and the pair with the smallest key can be found and removed quickly, and where the values can be updated. This implementation follows the idea of [http://docs.python.org/3.5/library/heapq.html](http://docs.python.org/3.5/library/heapq.html) where the updated elements are marked as removed. This means that the priority queue can be used unmodified. However, this might be expensive if changes are more common than popping (as might happen if the probability of choosing the best is close to zero).

In this implementation, the equal values are sorted randomly. This is achieved by having the elements of the heap being `[val, rand, elt]` triples, where the second element is a random number. Note that Python requires this to be a list, not a tuple, as the tuple cannot be modified.

cspSLS.py — (continued)

```python
class Updatable_priority_queue(object):
    """A priority queue where the values can be updated.
    Elements with the same value are ordered randomly.
    ""
    This code is based on the ideas described in
    http://docs.python.org/3.3/library/heapq.html
    It could probably be done more efficiently by
    shuffling the modified element in the heap.
    ""
    def __init__(self):
        self.pq = [] # priority queue of [val,rand,elt] triples
        self.elt_map = {} # map from elt to [val,rand,elt] triple in pq
        self.REMOVED = "*removed*" # a string that won't be a legal element
        self.max_size=0
    def add(self,elt,val):
        """adds elt to the priority queue with priority=val."
        assert val <= 0,val
        assert elt not in self.elt_map, elt
        new_triple = [val, random.random(),elt]
        heapq.heappush(self.pq, new_triple)
        self.elt_map[elt] = new_triple
```

http://aipython.org Version 0.9.4 March 2, 2022
def remove(self, elt):
    """remove the element from the priority queue""
    if elt in self.elt_map:
        self.elt_map[elt][2] = self.REMOVED
        del self.elt_map[elt]

def update_each_priority(self, update_dict):
    """update values in the priority queue by subtracting the values in
    update_dict from the priority of those elements in priority queue.
    ""
    for elt, incr in update_dict.items():
        if incr != 0:
            newval = self.elt_map.get(elt, [0])[0] - incr
            assert newval <= 0,
            str(elt)+"":"+(newval+incr)+"-"+str(incr)
            self.remove(elt)
        if newval != 0:
            self.add(elt, newval)

def pop(self):
    """Removes and returns the (elt,value) pair with minimal value.
    If the priority queue is empty, IndexError is raised.
    ""
    self.max_size = max(self.max_size, len(self.pq))  # keep statistics
    triple = heapq.heappop(self.pq)
    while triple[2] == self.REMOVED:
        triple = heapq.heappop(self.pq)
    del self.elt_map[triple[2]]
    return triple[2], triple[0]  # elt, value

def top(self):
    """Returns the (elt,value) pair with minimal value, without
    removing it.
    If the priority queue is empty, IndexError is raised.
    ""
    self.max_size = max(self.max_size, len(self.pq))  # keep statistics
    triple = self.pq[0]
    while triple[2] == self.REMOVED:
        heapq.heappop(self.pq)
    triple = self.pq[0]
    return triple[2], triple[0]  # elt, value

def empty(self):
    """returns True iff the priority queue is empty""
    return all(triple[2] == self.REMOVED for triple in self.pq)
4.5. Solving CSPs using Stochastic Local Search

4.5.4 Plotting Runtime Distributions

Runtime distribution uses matplotlib to plot runtime distributions. Here the runtime is a misnomer as we are only plotting the number of steps, not the time. Computing the runtime is non-trivial as many of the runs have a very short runtime. To compute the time accurately would require running the same code, with the same random seed, multiple times to get a good estimate of the runtime. This is left as an exercise.

cspSLS.py — (continued)

```python
import matplotlib.pyplot as plt
class Runtime_distribution(object):
    def __init__(self, csp, xscale='log'):
        """Sets up plotting for csp
        xscale is either 'linear' or 'log'
        """
        self.csp = csp
        plt.ion()
        plt.xlabel("Number of Steps")
        plt.ylabel("Cumulative Number of Runs")
        plt.xscale(xscale)  # Makes a 'log' or 'linear' scale

def plot_runs(self,num_runs=100,max_steps=1000, prob_best=1.0, prob_anycon=1.0):
    """Plots num_runs of SLS for the given settings.
    """
    stats = []
    SLSearcher.max_display_level, temp_mdl = 0,
    SLSearcher.max_display_level # no display
    for i in range(num_runs):
        searcher = SLSearcher(self.csp)
        num_steps = searcher.search(max_steps, prob_best, prob_anycon)
        if num_steps:
            stats.append(num_steps)
        stats.sort()
    if prob_best >= 1.0:
        label = "P(best)=1.0"
    else:
        p_ac = min(prob_anycon, 1-prob_best)
        label = "P(best)=%.2f, P(ac)=%.2f" % (prob_best, p_ac)
    plt.plot(stats,range(len(stats)),label=label)
    plt.legend(loc="upper left")
    SLSearcher.max_display_level= temp_mdl #restore display
```

4.5.5 Testing
4.6 Discrete Optimization

A SoftConstraint is a constraint, but where the condition is a real-valued function. Because we did not force the condition to be Boolean, we can make just reuse the Constraint class.

Exercise 4.17 Modify this to plot the runtime, instead of the number of steps. To measure runtime use timeit (https://docs.python.org/3.5/library/timeit.html). Small runtimes are inaccurate, so timeit can run the same code multiple times. Stochastic local algorithms give different runtimes each time called. To make the timing meaningful, you need to make sure the random seed is the same for each repeated call (see random.getstate and random.setstate in https://docs.python.org/3.5/library/random.html). Because the runtime for different seeds can vary a great deal, for each seed, you should start with 1 iteration and multiplying it by, say 10, until the time is greater than 0.2 seconds. Make sure you plot the average time for each run. Before you start, try to estimate the total runtime, so you will be able to tell if there is a problem with the algorithm stopping.
class SoftConstraint(Constraint):
    """A Constraint consists of
    * scope: a tuple of variables
    * function: a real-valued function that can applied to a tuple of values
    * string: a string for printing the constraints. All of the strings
      must be unique.
    for the variables
    """
    def __init__(self, scope, function, string=None, position=None):
        Constraint.__init__(self, scope, function, string, position)

    def value(self, assignment):
        return self.holds(assignment)

cspSoft.py — (continued)

A = Variable('A', {1,2}, position=(0.2,0.9))
B = Variable('B', {1,2,3}, position=(0.8,0.9))
C = Variable('C', {1,2}, position=(0.5,0.5))
D = Variable('D', {1,2}, position=(0.8,0.1))

def c1fun(a,b):
    if a==1:
        return (5 if b==1 else 2)
    else:
        return (0 if b==1 else 4 if b==2 else 3)
c1 = SoftConstraint([A,B],c1fun,"c1")
def c2fun(b,c):
    if b==1:
        return (5 if c==1 else 2)
    elif b==2:
        return (0 if c==1 else 4)
    else:
        return (2 if c==1 else 0)
c2 = SoftConstraint([B,C],c2fun,"c2")
def c3fun(b,d):
    if b==1:
        return (3 if d==1 else 0)
    elif b==2:
        return 2
    else:
        return (2 if d==1 else 4)
c3 = SoftConstraint([B,D],c3fun,"c3")
def penalty_if_same(pen):
    """returns a function that gives a penalty of pen if the arguments are
    the same"
    return lambda x,y: (pen if (x==y) else 0)
c4 = SoftConstraint([C,A],penalty_if_same(3),"c4")
scsp1 = CSP("scsp1", [A,B,C,D], [c1,c2,c3,c4])

### The second soft CSP has an extra variable, and 2 constraints
E = Variable('E', {1,2}, position=(0.1,0.1))
c5 = SoftConstraint([C,E],penalty_if_same(3),"c5")
c6 = SoftConstraint([D,E],penalty_if_same(2),"c6")
scsp2 = CSP("scsp1", [A,B,C,D,E], [c1,c2,c3,c4,c5,c6])
4.6.1 Branch-and-bound Search

Here we specialize the branch-and-bound algorithm (Section 3.3 on page 47).

```python
from display import Displayable, visualize
import math

class DF_branch_and_bound_opt(Displayable):
    """returns a branch and bound search a problem.
    An optimal assignment with cost less than bound can be found by calling
    search()
    """
    def __init__(self, csp, bound=math.inf):
        """creates a searcher than can be used with search() to find an
        optimal path.
        bound gives the initial bound. By default this is infinite -
        meaning there
        is no initial pruning due to depth bound
        """
        super().__init__()
        self.csp = csp
        self.best_asst = None
        self.bound = bound

    def optimize(self):
        """returns an optimal solution to a problem with cost less than
        bound.
        returns None if there is no solution with cost less than bound."""
        self.num_expanded=0
        self.cbsearch({}, 0, self.csp.constraints)
        self.display(1,"Number of paths expanded:",self.num_expanded)
        return self.best_asst, self.bound

    def cbsearch(self, asst, cost, constraints):
        """finds the optimal solution that extends path and is less the
        bound"
        self.display(2,"cbsearch:",asst,cost,constraints)
        can_eval = [c for c in constraints if c.can_evaluate(asst)]
        rem_cons = [c for c in constraints if c not in can_eval]
        newcost = cost + sum(c.value(asst) for c in can_eval)
        self.display(2,"Evaluaing:",can_eval,"cost:",newcost)
        if newcost < self.bound:
            self.num_expanded += 1
            if rem_cons==[]:
                self.best_asst = asst
                self.bound = newcost
                self.display(1,"New best assignment:",asst," cost:",newcost)
            else:
                var = next(var for var in self.csp.variables if var not in asst)
```

http://aipython.org  Version 0.9.4  March 2, 2022
for val in var.domain:
    self.cbsearch({var:val}|asst, newcost, rem_cons)

# bnb = DF_branch_and_bound_opt(scsp1)
# bnb.max_display_level=3 # show more detail
# bnb.optimize()
## 5.1 Representing Knowledge Bases

A clause consists of a head (an atom) and a body. A body is represented as a list of atoms. Atoms are represented as strings.

```python
class Clause(object):
    """A definite clause""
    def __init__(self, head, body=[]):
        """clause with atom head and lost of atoms body""
        self.head = head
        self.body = body

    def __str__(self):
        """returns the string representation of a clause."
        if self.body:
            return self.head + " <- " + " & ".join(self.body) + "."
        else:
            return self.head + "."
```

An askable atom can be asked of the user. The user can respond in English or French or just with a “y”.

```python
class Askable(object):
    """An askable atom""
    def __init__(self, atom):
        """clause with atom head and lost of atoms body""
```
A knowledge base is a list of clauses and askables. In order to make top-down inference faster, this creates a dictionary that maps each atom into the set of clauses with that atom in the head.
5.1. Representing Knowledge Bases

Here is a representation of the electrical domain of the textbook:

```python
elect = KB(
    Clause('light_l1'),
    Clause('light_l2'),
    Clause('ok_l1'),
    Clause('ok_l2'),
    Clause('ok_cb1'),
    Clause('ok_cb2'),
    Clause('live_outside'),
    Clause('live_l1', ['live_w0']),
    Clause('live_w0', ['up_s2', 'live_w1']),
    Clause('live_w0', ['down_s2', 'live_w2']),
    Clause('live_w1', ['up_s1', 'live_w3']),
    Clause('live_w2', ['down_s1', 'live_w3']),
    Clause('live_l2', ['live_w4']),
    Clause('live_w4', ['up_s3', 'live_w3']),
    Clause('live_p_1', ['live_w3']),
    Clause('live_w3', ['live_w5', 'ok_cb1']),
    Clause('live_p_2', ['live_w6']),
    Clause('live_w6', ['live_w5', 'ok_cb2']),
    Clause('live_w5', ['live_outside']),
    Clause('lit_l1', ['light_l1', 'live_l1', 'ok_l1']),
    Clause('lit_l2', ['light_l2', 'live_l2', 'ok_l2']),
    Askable('up_s1'),
    Askable('down_s1'),
    Askable('up_s2'),
    Askable('down_s2'),
    Askable('up_s3'),
    Askable('down_s2')
)

# print(kb)
```

The following knowledge base is false of the intended interpretation. One of the clauses is wrong; can you see which one? We will show how to debug it.

```python
elect_bug = KB(
    Clause('light_l2'),
    Clause('ok_l1'),
    Clause('ok_l2'),
    Clause('ok_cb1'),
    Clause('ok_cb2'),
    Clause('live_outside'),
    Clause('live_p_2', ['live_w6']),
```
5. Propositions and Inference

5.2 Bottom-up Proofs (with askables)

`fixed_point` computes the fixed point of the knowledge base `kb`.
5.2. Bottom-up Proofs (with askables)

```python
# 
fp = ask_askables(kb)
added = True
while added:
    added = False # added is true when an atom was added to fp this iteration
    for c in kb.clauses:
        if c.head not in fp and all(b in fp for b in c.body):
            fp.add(c.head)
            added = True
        kb.display(2,c.head,"added to fp due to clause",c)
return fp

def ask_askables(kb):
    return {at for at in kb.askables if yes(input("Is \+at\+ true? "))}
```

The following provides a trivial unit test, by default using the knowledge base triv_KB:

```python
from logicProblem import triv_KB
def test(kb=triv_KB, fixedpt = {'i_am','i_think'}):
    fp = fixed_point(kb)
    assert fp == fixedpt, "kb gave result "+str(fp)
print("Passed unit test")
if __name__ == "__main__":
    test()

from logicProblem import elect
# elect.max_display_level=3 # give detailed trace
# fixed_point(elect)
```

Exercise 5.1 It is not very user-friendly to ask all of the askables up-front. Implement ask-the-user so that questions are only asked if useful, and are not re-asked. For example, if there is a clause $h \leftarrow a \land b \land c \land d \land e$, where $c$ and $e$ are askable, $c$ and $e$ only need to be asked if $a, b, d$ are all in $fp$ and they have not been asked before. Askable $e$ only needs to be asked if the user says “yes” to $c$. Askable $c$ doesn’t need to be asked if the user previously replied “no” to $e$.

This form of ask-the-user can ask a different set of questions than the top-down interpreter that asks questions when encountered. Give an example where they ask different questions (neither set of questions asked is a subset of the other).

Exercise 5.2 This algorithm runs in time $O(n^2)$, where $n$ is the number of clauses, for a bounded number of elements in the body; each iteration goes through each of the clauses, and in the worst case, it will do an iteration for each clause. It is possible to implement this in time $O(n)$ time by creating an index that maps an atom to the set of clauses with that atom in the body. Implement this. What is its complexity as a function of $n$ and $b$, the maximum number of atoms in the body of a clause?

http://aipython.org

Version 0.9.4 March 2, 2022
Exercise 5.3  It is possible to be asymptotically more efficient (in terms of the number of elements in a body) than the method in the previous question by noticing that each element of the body of clause only needs to be checked once. For example, the clause \( a \leftarrow b \land c \land d \), needs only be considered when \( b \) is added to \( fp \). Once \( b \) is added to \( fp \), if \( c \) is already in \( pf \), we know that \( a \) can be added as soon as \( d \) is added. Implement this. What is its complexity as a function of \( n \) and \( b \), the maximum number of atoms in the body of a clause?

5.3 Top-down Proofs (with askables)

prove\((kb, goal)\) is used to prove goal from a knowledge base, \( kb \), where a goal is a list of atoms. It returns True if \( kb \vdash goal \). The indent is used when displaying the code (and doesn’t need to have a non-default value).

```python
from logicProblem import yes
def prove(kb, ans_body, indent=" ":
    """returns True if kb |- ans_body
    ans_body is a list of atoms to be proved
    ""
    kb.display(2,indent,'yes <-',' & '.join(ans_body))
    if ans_body:
        selected = ans_body[0] # select first atom from ans_body
        if selected in kb.askables:
            return (yes(input("Is "+selected+" true? "))
            and prove(kb,ans_body[1:],indent+" "))
        else:
            return any(prove(kb,cl.body+ans_body[1:],indent+" ")
            for cl in kb.clauses_for_atom(selected))
    else:
        return True # empty body is true
```

The following provides a simple unit test that is hard wired for triv_KB:

```python
from logicProblem import triv_KB
def test():
    a1 = prove(triv_KB,['i_am'])
    assert a1, "triv_KB proving i_am gave "+str(a1)
    a2 = prove(triv_KB,['i_smell'])
    assert not a2, "triv_KB proving i_smell gave "+str(a2)
    print("Passed unit tests")
if __name__ == "__main__":
    test()
```

http://aipython.org  Version 0.9.4  March 2, 2022
Exercise 5.4 This code can re-ask a question multiple times. Implement this code so that it only asks a question once and remembers the answer. Also implement a function to forget the answers.

Exercise 5.5 What search method is this using? Implement the search interface so that it can use $A^*$ or other searching methods. Define an admissible heuristic that is not always 0.

5.4 Debugging and Explanation

Here we modify the top-down procedure to build a proof tree than can be traversed for explanation and debugging.

prove_atom(kb, atom) returns a proof for atom from a knowledge base kb, where a proof is a pair of the atom and the proofs for the elements of the body of the clause used to prove the atom. prove_body(kb, body) returns a list of proofs for list body from a knowledge base, kb. The indent is used when displaying the code (and doesn’t need to have a non-default value).

```python
from logicProblem import yes # for asking the user

def prove_atom(kb, atom, indent=""): 
    """returns a pair (atom,proofs) where proofs is the list of proofs
    of the elements of a body of a clause used to prove atom.
    """
    kb.display(2,indent,'proving',atom)
    if atom in kb.askables:
        if yes(input("Is "+atom+" true? ")):
            return (atom,"answered")
        else:
            return "fail"
    else:
        for cl in kb.clauses_for_atom(atom):
            kb.display(2,indent,"trying",atom,"<-", & .join(cl.body))
            pr_body = prove_body(kb, cl.body, indent)
            if pr_body != "fail":
                return (atom, pr_body)
        return "fail"

def prove_body(kb, ans_body, indent=""): 
    """returns proof tree if kb |- ans_body or "fail" if there is no proof
    ans_body is a list of atoms in a body to be proved
    """
    proofs = []
    for atom in ans_body:
        proof_at = prove_atom(kb, atom, indent+" ")
        if proof_at == "fail":
            return "fail" # fail if any proof fails
        else:
```

http://aipython.org
The following provides a simple unit test that is hard wired for \texttt{triv\_KB}:

```python
from logicProblem import triv_KB
def test():
a1 = prove_atom(triv_KB,'i_am')
assert a1, "triv\_KB proving i_am gave "+str(a1)
a2 = prove_atom(triv_KB,'i_smell')
assert a2=="fail", "triv\_KB proving i_smell gave "+str(a2)
print("Passed unit tests")
if __name__ == "__main__":
test()
```

The \texttt{interact(kb)} provides an interactive interface to explore proofs for knowledge base \texttt{kb}. The user can ask to prove atoms and can ask how an atom was proved.

To ask how, there must be a current atom for which there is a proof. This starts as the atom asked. When the user asks “how n” the current atom becomes the n-th element of the body of the clause used to prove the (previous) current atom. The command “up” makes the current atom the atom in the head of the rule containing the (previous) current atom. Thus “how n” moves down the proof tree and “up” moves up the proof tree, allowing the user to explore the full proof.

```python
helptext = """Commands are:

\begin{itemize}
\item \texttt{ask atom} \hspace{1cm} ask is there is a proof for atom (atom should not be in quotes)
\item \texttt{how} \hspace{1cm} show the clause that was used to prove atom
\item \texttt{how n} \hspace{1cm} show the clause used to prove the nth element of the body
\item \texttt{up} \hspace{1cm} go back up proof tree to explore other parts of the proof tree
\item \texttt{kb} \hspace{1cm} print the knowledge base
\item \texttt{quit} \hspace{1cm} quit this interaction (and go back to Python)
\item \texttt{help} \hspace{1cm} print this text
\end{itemize}

\""

def interact(kb):
    going = True
    ups = [] # stack for going up
    proof="fail" # there is no proof to start
    while going:
        inp = input("logicExplain: ")
        inps = inp.split(" ")
```

\url{http://aipython.org}
try:
    command = inps[0]
    if command == "quit":
        going = False
    elif command == "ask":
        proof = prove_atom(kb, inps[1])
        if proof == "fail":
            print("fail")
        else:
            print("yes")
    elif command == "how":
        if proof == "fail":
            print("there is no proof")
        elif len(inps) == 1:
            print_rule(proof)
        else:
            try:
                ups.append(proof)
                proof = proof[1][int(inps[1])] # nth argument of rule
                print_rule(proof)
            except:
                print("In "how n", n must be a number between 0 and", len(proof[1])-1,"inclusive.")
    elif command == "up":
        if ups:
            proof = ups.pop()
        else:
            print("No rule to go up to.")
            print_rule(proof)
    elif command == "kb":
        print(kb)
    elif command == "help":
        print(helptext)
    else:
        print("unknown command:", inp)
        print("use help for help")
except:
    print("unknown command:", inp)
    print("use help for help")

def print_rule(proof):
    (head, body) = proof
    if body == "answered":
        print(head, "was answered yes")
    elif body == []:
        print(head, "is a fact")
    else:
        print(head, "<-")
        for i, a in enumerate(body):
            print(i, ":" , a[0])
The following shows an interaction for the knowledge base `elect`:

```python
>>> interact(elect)
logicExplain: ask lit_l1
Is up_s2 true? no
Is down_s2 true? yes
Is down_s1 true? yes
   logicExplain: how
lit_l1 <-
   0 : light_l1
   1 : live_l1
   2 : ok_l1
logicExplain: how 1
live_l1 <-
   0 : live_w0
logicExplain: how 0
live_w0 <-
   0 : down_s2
   1 : live_w2
logicExplain: how 0
down_s2 was answered yes
logicExplain: up
live_w0 <-
   0 : down_s2
   1 : live_w2
logicExplain: how 1
live_w2 <-
   0 : down_s1
   1 : live_w3
logicExplain: quit
```
5.5 Assumables

Atom \( a \) can be made assumable by including \( \text{Assumable}(a) \) in the knowledge base. A knowledge base that can include assumables is declared with \( \text{KBA} \).

```python
from logicProblem import Clause, Askable, KB, yes
class Assumable(object):
    """An askable atom""
    def __init__(self, atom):
        """clause with atom head and lost of atoms body""
        self.atom = atom
    def __str__(self):
        """returns the string representation of a clause."
        return "assumable " + self.atom + "."
class KBA(KB):
    """A knowledge base that can include assumables""
    def __init__(self, statements):
        self.assumables = [c.atom for c in statements if isinstance(c, Assumable)]
        KB.__init__(self, statements)

The top-down Horn clause interpreter, \( \text{prove\_all\_ass} \) returns a list of the sets of assumables that imply \( \text{ans\_body} \). This list will contain all of the minimal sets of assumables, but can also find non-minimal sets, and repeated sets, if they can be generated with separate proofs. The set \( \text{assumed} \) is the set of assumables already assumed.

```
Given a list of sets, $\text{minsets}$ returns a list of the minimal sets in the list. For example, $\text{minsets}([\{2,3,4\}, \{2,3\}, \{6,2,3\}, \{2,3\}, \{2,4,5\}])$ returns $[\{2,3\}, \{2,4,5\}]$.

```
def minsets(ls):
    """ls is a list of sets
    returns a list of minimal sets in ls
    ""
    ans = []  # elements known to be minimal
    for c in ls:
        if not any(c1 < c for c1 in ls) and not any(c1 <= c for c1 in ans):
            ans.append(c)
    return ans
```

Warning: $\text{minsets}$ works for a list of sets or for a set of (frozen) sets, but it does not work for a generator of sets. For example, try to predict and then test:

```
minsets(e for e in [[2, 3, 4], [2, 3], [6, 2, 3], [2, 3], [2, 4, 5]])
```

The diagnoses can be constructed from the (minimal) conflicts as follows. This also works if there are non-minimal conflicts, but is not as efficient.

```
def diagnoses(cons):
    """cons is a list of (minimal) conflicts.
    returns a list of diagnoses."
    if cons == []:
        return [set()]
    else:
        return minsets([[e]|d)  # | is set union
                        for e in cons[0]
                        for d in diagnoses(cons[1:]))
```

Test cases:

```
electa = KBA([  
    Clause('light_l1'),
])
```
5.5. Assumables

Clause('light_l2'),
Assumable('ok_l1'),
Assumable('ok_l2'),
Assumable('ok_s1'),
Assumable('ok_s2'),
Assumable('ok_s3'),
Assumable('ok_cb1'),
Assumable('ok_cb2'),
Assumable('live_outside'),
Clause('live_l1', ['live_w0']),
Clause('live_w0', ['up_s2', 'ok_s2', 'live_w1']),
Clause('live_w0', ['down_s2', 'ok_s2', 'live_w2']),
Clause('live_w1', ['up_s1', 'ok_s1', 'live_w3']),
Clause('live_w2', ['down_s1', 'ok_s1', 'live_w3']),
Clause('live_l2', ['live_w4']),
Clause('live_w4', ['up_s3', 'ok_s3', 'live_w3']),
Clause('live_p_1', ['live_w3']),
Clause('live_w3', ['live_w5', 'ok_cb1']),
Clause('live_p_2', ['live_w6']),
Clause('live_w6', ['live_w5', 'ok_cb2']),
Clause('live_w5', ['live_w4', 'live_outside']),
Clause('lit_l1', ['light_l1', 'live_l1', 'ok_l1']),
Clause('lit_l2', ['light_l2', 'live_l2', 'ok_l2']),
Askable('up_s1'),
Askable('down_s1'),
Askable('up_s2'),
Askable('down_s2'),
Askable('up_s3'),
Askable('down_s3'),
Askable('dark_l1'),
Askable('dark_l2'),
Clause('false', ['dark_l1', 'lit_l1']),
Clause('false', ['dark_l2', 'lit_l2'])
)
# electa.prove_all_ass(['false'])
# cs=electa.conflicts()
# print(cs)
# diagnoses(cs)  # diagnoses from conflicts

Exercise 5.7 To implement a version of conflicts that never generates non-minimal conflicts, modify prove_all_ass to implement iterative deepening on the number of assumables used in a proof, and prune any set of assumables that is a superset of a conflict.

Exercise 5.8 Implement explanations(self, body), where body is a list of atoms, that returns the a list of the minimal explanations of the body. This does not require modification of prove_all_ass.

Exercise 5.9 Implement explanations, as in the previous question, so that it never generates non-minimal explanations. Hint: modify prove_all_ass to implement iter-
ative deepening on the number of assumptions, generating conflicts and explanations together, and pruning as early as possible.
Chapter 6

Planning with Certainty

6.1 Representing Actions and Planning Problems

The STRIPS representation of an action consists of:

- the name of the action
- preconditions: a dictionary of feature:value pairs that specifies that the feature must have this value for the action to be possible
- effects: a dictionary of feature:value pairs that are made true by this action. In particular, a feature in the dictionary has the corresponding value (and not its previous value) after the action, and a feature not in the dictionary keeps its old value.

class Strips(object):
    def __init__(self, name, preconds, effects, cost=1):
        
        defines the STRIPS representation for an action:
        * name is the name of the action
        * preconds, the preconditions, is feature:value dictionary that
          must hold
        for the action to be carried out
        * effects is a feature:value map that this action makes
          true. The action changes the value of any feature specified
          here, and leaves other features unchanged.
        * cost is the cost of the action

        

101
A STRIPS domain consists of:

- A set of actions.
- A dictionary that maps each feature into a set of possible values for the feature.
- A list of the actions

A planning problem consists of a planning domain, an initial state, and a goal. The goal does not need to fully specify the final state.

### 6.1.1 Robot Delivery Domain

The following specifies the robot delivery domain of Section 6.1, shown in Figure [6.1](http://aipython.org).

```python
class STRIPS_domain(object):
    def __init__(self, feature_domain_dict, actions):
        """Problem domain
        feature_domain_dict is a feature:domain dictionary,
        mapping each feature to its domain
        actions
        ""
        self.feature_domain_dict = feature_domain_dict
        self.actions = actions
```

```python
class Planning_problem(object):
    def __init__(self, prob_domain, initial_state, goal):
        """a planning problem consists of
        * a planning domain
        * the initial state
        * a goal
        ""
        self.prob_domain = prob_domain
        self.initial_state = initial_state
        self.goal = goal
```
### 6.1. Representing Actions and Planning Problems

**Features to describe states**

- **RLoc** – Rob’s location
- **RHC** – Rob has coffee
- **SWC** – Sam wants coffee
- **MW** – Mail is waiting
- **RHM** – Rob has mail

**Actions**

- **mc** – move clockwise
- **mcc** – move counterclockwise
- **puc** – pickup coffee
- **dc** – deliver coffee
- **pum** – pickup mail
- **dm** – deliver mail

---

```python
boolean = {True, False}
delivery_domain = STRIPS_domain(
    {'RLoc': 'cs', 'off', 'lab', 'mr'}, 'RHC': boolean, 'SWC': boolean,
    'MW': boolean, 'RHM': boolean), # feature: value dictionary
    { Strips('mc_cs', {'RLoc': 'cs'}, {'RLoc': 'off'}),
      Strips('mc_off', {'RLoc': 'off'}, {'RLoc': 'lab'}),
      Strips('mc_lab', {'RLoc': 'lab'}, {'RLoc': 'mr'}),
      Strips('mc_mr', {'RLoc': 'mr'}, {'RLoc': 'cs'}),
      Strips('mcc_cs', {'RLoc': 'cs'}, {'RLoc': 'mr'}),
      Strips('mcc_off', {'RLoc': 'off'}, {'RLoc': 'cs'}),
      Strips('mcc_lab', {'RLoc': 'lab'}, {'RLoc': 'off'}),
      Strips('mcc_mr', {'RLoc': 'mr'}, {'RLoc': 'lab'}),
      Strips('puc', {'RLoc': 'cs', 'RHC': False}, {'RHC': True}),
      Strips('dc', {'RLoc': 'off', 'RHC': True}, {'RHC': False, 'SWC': False}),
      Strips('pum', {'RLoc': 'mr', 'MW': True}, {'RHM': True, 'MW': False}),
      Strips('dm', {'RLoc': 'off', 'RHM': True}, {'RHM': False})
    })
```

---

http://aipython.org  Version 0.9.4  March 2, 2022
6. Planning with Certainty

6.1.2 Blocks World

The blocks world consist of blocks and a table. Each block can be on the table or on another block. A block can only have one other block on top of it. Figure 6.2 shows 3 states with some of the actions between them.

A state is defined by the two features:

- **on** where on(x) = y when block x is on block or table y
- **clear** where clear(x) = True when block x has nothing on it.

There is one parameterized action

- move(x, y, z) move block x from y to z, where y and z could be a block or the table.

```python
problem0 = Planning_problem(delivery_domain,
    {'RLoc':'lab', 'MW':True, 'SWC':True, 'RHC':False, 'RHM':False})
problem1 = Planning_problem(delivery_domain,
    {'RLoc':'lab', 'MW':True, 'SWC':True, 'RHC':False, 'RHM':False})
problem2 = Planning_problem(delivery_domain,
    {'RLoc':'lab', 'MW':True, 'SWC':True, 'RHC':False, 'RHM':False})
```

Figure 6.2: Blocks world with two actions
To handle parameterized actions (which depend on the blocks involved), the actions and the features are all strings, created for the all combinations of the blocks. Note that we treat moving to a block separately from moving to the table, because the blocks needs to be clear, but the table always has room for another block.

```python
## blocks world

```move(x,y,z):
    """string for the 'move' action""
    return 'move_'+x+'_from_']+y_to_'] z
```

```python
don(x):
    """string for the 'on' feature""
    return x'+_is_on
```

```python
clear(x):
    """string for the 'clear' feature""
    return 'clear_'+x
```

```python
def create_blocks_world(blocks = {'a', 'b', 'c', 'd'}):
    blocks_and_table = blocks | {'table'}
    stmap = {Strips(move(x,y,z),{on(x):y, clear(x):True, clear(z):True},
    {on(x):z, clear(y):True, clear(z):False})
    for x in blocks
    for y in blocks_and_table
    for z in blocks
    if x!=y and y!=z and z!=x}
    stmap.update({Strips(move(x,y,'table'), {on(x):y, clear(x):True},
    {on(x):'table', clear(y):True})
    for x in blocks
    for y in blocks
    if x!=y})
    feature_domain_dict = {on(x):blocks_and_table-{x} for x in blocks}
    feature_domain_dict.update({clear(x):boolean for x in blocks_and_table})
    return STRIPS_domain(feature_domain_dict, stmap)
```

The problem blocks1 is a classic example, with 3 blocks, and the goal consists of two conditions. See Figure 6.3. Note that this example is challenging because we can’t achieve one of the goals and then the other; whichever one we achieve first has to be undone to achieve the second.

```python
blocks1dom = create_blocks_world({'a', 'b', 'c'})
blocks1 = Planning_problem(blocks1dom,
{on('a'):'table', clear('a'):True,
 on('b'):'c', clear('b'):True,
 on('c'):'table', clear('c'):False}, # initial state
{on('a'):'b', on('c'):'a'}) #goal
```

The problem blocks2 is one to invert a tower of size 4.

```python
blocks2dom = create_blocks_world({'a', 'b', 'c', 'd'})
```
6. Planning with Certainty

Figure 6.3: Blocks problem blocks1

```python
119 tower4 = {clear('a'):True, on('a'): 'b',
120     clear('b'):False, on('b'): 'c',
121     clear('c'):False, on('c'): 'd',
122     clear('d'):False, on('d'): 'table'}
123 blocks2 = Planning_problem(blocks2dom,
124     tower4, # initial state
125     {on('d'): 'c', on('c'): 'b', on('b'): 'a')} #goal

The problem blocks3 is to move the bottom block to the top of a tower of size 4.

```stripsProblem.py — (continued)```

```python
127 blocks3 = Planning_problem(blocks2dom,
128     tower4, # initial state
129     {on('d'): 'a', on('a'): 'b', on('b'): 'c')} #goal

Exercise 6.1 Represent the problem of given a tower of 4 blocks (a on b on c on d on table), the goal is to have a tower with the previous top block on the bottom (b on c on d on a). Do not include the table in your goal (the goal does not care whether a is on the table). [Before you run the program, estimate how many steps it will take to solve this.] How many steps does an optimal planner take?

Exercise 6.2 Represent the domain so that on(x, y) is a Boolean feature that is True when x is on y, Does the representation of the state need to not include negative on facts? Why or why not? (Note that this may depend on the planner; write your answer with respect to particular planners.)

Exercise 6.3 It is possible to write the representation of the problem without using clear, where clear(x) means nothing is on x. Change the definition of the blocks world so that it does not use clear but uses on being false instead. Does this work better for any of the planners?

6.2 Forward Planning

To run the demo, in folder "aipython", load "stripsForwardPlanner.py", and copy and paste the commented-out example queries at the bottom of that file.

http://aipython.org  Version 0.9.4  March 2, 2022
6.2. Forward Planning

In a forward planner, a node is a state. A state consists of an assignment, which is a variable:value dictionary. In order to be able to do multiple-path pruning, we need to define a hash function, and equality between states.

```python
from searchProblem import Arc, Search_problem
from stripsProblem import Strips, STRIPS_domain

class State(object):
    def __init__(self, assignment):
        self.assignment = assignment
        self.hash_value = None
    def __hash__(self):
        if self.hash_value is None:
            self.hash_value = hash(frozenset(self.assignment.items()))
        return self.hash_value
    def __eq__(self, st):
        return self.assignment == st.assignment
    def __str__(self):
        return str(self.assignment)
```

In order to define a search problem (page 33), we need to define the goal condition, the start nodes, the neighbours, and (optionally) a heuristic function. Here zero is the default heuristic function.

```python
def zero(*args,**nargs):
    """always returns 0"""
    return 0

class Forward_STRIPS(Search_problem):
    """A search problem from a planning problem where:
    * a node is a state object.
    * the dynamics are specified by the STRIPS representation of actions
    """
    def __init__(self, planning_problem, heur=zero):
        """creates a forward search space from a planning problem.
        heur(state,goal) is a heuristic function,
        an underestimate of the cost from state to goal, where
        both state and goals are feature:value dictionaries.
        """
        self.prob_domain = planning_problem.prob_domain
        self.initial_state = State(planning_problem.initial_state)
        self.goal = planning_problem.goal
        self.heur = heur

    def is_goal(self, state):
        """is True if node is a goal.
        Every goal feature has the same value in the state and the goal."""
        return all(state.assignment[prop]==self.goal[prop]
```
for prop in self.goal)

def start_node(self):
    """returns start node""
    return self.initial_state

def neighbors(self,state):
    """returns neighbors of state in this problem""
    return [ Arc(state, self.effect(act,state.assignment), act.cost, act)
               for act in self.prob_domain.actions
               if self.possible(act,state.assignment)]

def possible(self,act,state_asst):
    """True if act is possible in state.
    act is possible if all of its preconditions have the same value in
    the state""
    return all(state_asst[pre] == act.preconds[pre]
                for pre in act.preconds)

def effect(self,act,state_asst):
    """returns the state that is the effect of doing act given
    state_asst
    Python 3.9: return state_asst | act.effects"
    new_state_asst = state_asst.copy()
    new_state_asst.update(act.effects)
    return State(new_state_asst)

def heuristic(self,state):
    """in the forward planner a node is a state.
    the heuristic is an (under)estimate of the cost
    of going from the state to the top-level goal.
    ""
    return self.heur(state.assignment, self.goal)

Here are some test cases to try.

# SearcherMPP(Forward_STRIPS(problem1)).search() #A* with MPP
# DF_branch_and_bound(Forward_STRIPS(problem1),10).search() #B&B
# To find more than one plan:
# s1 = SearcherMPP(Forward_STRIPS(problem1)) #A*
# s1.search() #find another plan

http://aipython.org  Version 0.9.4  March 2, 2022
6.2. Forward Planning

6.2.1 Defining Heuristics for a Planner

Each planning domain requires its own heuristics. If you change the actions, you will need to reconsider the heuristic function, as there might then be a lower-cost path, which might make the heuristic non-admissible.

Here is an example of defining a (not very good) heuristic for the coffee delivery planning domain.

First we define the distance between two locations, which is used for the heuristics.

```python
def dist(loc1, loc2):
    """returns the distance from location loc1 to loc2
    ""
    if loc1==loc2:
        return 0
    if {loc1,loc2} in [{'cs','lab'},{'mr','off'}]:
        return 2
    else:
        return 1
```

Note that the current state is a complete description; there is a value for every feature. However the goal need not be complete; it does not need to define a value for every feature. Before checking the value for a feature in the goal, a heuristic needs to define whether the feature is defined in the goal.

```python
def h1(state,goal):
    """ the distance to the goal location, if there is one"
    if 'RLoc' in goal:
        return dist(state['RLoc'], goal['RLoc'])
    else:
        return 0
```

```python
def h2(state,goal):
    """ the distance to the coffee shop plus getting coffee and delivering it
    if the robot needs to get coffee
    ""
    if ('SWC' in goal and goal['SWC']==False
        and state['SWC']==True
        and state['RHC']==False):
        return dist(state['RLoc'],'cs')+3
    else:
        return 0
```

The maximum of the values of a set of admissible heuristics is also an admissible heuristic. The function `maxh` takes a number of heuristic functions as arguments, and returns a new heuristic function that takes the maximum of the values of the heuristics. For example, `h1` and `h2` are heuristic functions and so `maxh(h1,h2)` is also. `maxh` can take an arbitrary number of arguments.
def maxh(*heuristics):
    """Returns a new heuristic function that is the maximum of the
    functions in heuristics. 
    heuristics is the list of arguments which must be heuristic functions.
    """
    # return lambda state,goal: max(h(state,goal) for h in heuristics)
    def newh(state,goal):
        return max(h(state,goal) for h in heuristics)
    return newh

The following runs the example with and without the heuristic.

def test_forward_heuristic(thisproblem=problem1):
    print("\n***** FORWARD NO HEURISTIC")
    print(SearcherMPP(Forward_STRIPS(thisproblem)).search())
    print("\n***** FORWARD WITH HEURISTIC h1")
    print(SearcherMPP(Forward_STRIPS(thisproblem,h1)).search())
    print("\n***** FORWARD WITH HEURISTIC h2")
    print(SearcherMPP(Forward_STRIPS(thisproblem,h2)).search())
    print("\n***** FORWARD WITH HEURISTICs h1 and h2")
    print(SearcherMPP(Forward_STRIPS(thisproblem,maxh(h1,h2))).search())

if __name__ == "__main__":
    test_forward_heuristic()

Exercise 6.4 Try the forward planner with a heuristic function of just $h_1$, with just $h_2$ and with both. Explain how each one prunes or doesn’t prune the search space.

Exercise 6.5 Create a better heuristic than $\text{maxh}(h_1,h_2)$. Try it for a number of different problems. In particular, try and include the following costs:

i) $h_3$ is like $h_2$ but also takes into account the case when $Rloc$ is in goal.

ii) $h_4$ uses the distance to the mail room plus getting mail and delivering it if the robot needs to get need to deliver mail.

iii) $h_5$ is for getting mail when goal is for the robot to have mail, and then getting to the goal destination (if there is one).

Exercise 6.6 Create an admissible heuristic for the blocks world.
6.3 Regression Planning

To run the demo, in folder "aipython", load "stripsRegressionPlanner.py", and copy and paste the commented-out example queries at the bottom of that file.

In a regression planner a node is a subgoal that need to be achieved.

A Subgoal object consists of an assignment, which is variable:value dictionary. We make it hashable so that multiple path pruning can work. The hash is only computed when necessary (and only once).

A regression search has subgoals as nodes. The initial node is the top-level goal of the planner. The goal for the search (when the search can stop) is a subgoal that holds in the initial state.
6. Planning with Certainty

```python
# def is_goal(self, subgoal):
    # """if subgoal is true in the initial state, a path has been found""
    # goal_asst = subgoal.assignment
    # return all(self.initial_state[g]==goal_asst[g]
    #     for g in goal_asst)

# def start_node(self):
    # """the start node is the top-level goal""
    # return self.top_goal

# def neighbors(self, subgoal):
    # """returns a list of the arcs for the neighbors of subgoal in this problem""
    # goal_asst = subgoal.assignment
    # return [
    #     Arc(subgoal, self.weakest_precond(act,goal_asst),
    #         act.cost, act)
    #     for act in self.prob_domain.actions
    #     if self.possible(act,goal_asst)]

# def possible(self, act, goal_asst):
    # """True if act is possible to achieve goal_asst.
    # the action achieves an element of the effects and
    # the action doesn't delete something that needs to be achieved and
    # the preconditions are consistent with other subgoals that need to be achieved
    # ""
    # return
    #     any(goal_asst[prop] == act.effects[prop]
    #         for prop in act.effects if prop in goal_asst)
    #     and all(goal_asst[prop] == act.effects[prop]
    #         for prop in act.effects if prop in goal_asst)
    #     and all(goal_asst[prop]== act.preconds[prop]
    #         for prop in act.preconds if prop not in act.effects
    #             and prop in goal_asst)

# def weakest_precond(self, act, goal_asst):
    # """returns the subgoal that must be true so goal_asst holds after act
    # should be: act.preconds | (goal_asst - act.effects)
    # ""
    # new_asst = act.preconds.copy()
    # for g in goal_asst:
    #     if g not in act.effects:
    #         new_asst[g] = goal_asst[g]
    # return Subgoal(new_asst)

# def heuristic(self, subgoal):
    # """in the regression planner a node is a subgoal.
```
the heuristic is an (under)estimate of the cost of going from the initial state to subgoal.

"""
return self.heur(self.initial_state, subgoal.assignment)
"""

6.3.1 Defining Heuristics for a Regression Planner

The regression planner can use the same heuristic function as the forward planner. However, just because a heuristic is useful for a forward planner does not mean it is useful for a regression planner, and vice versa. You should experiment with whether the same heuristic works well for both a regression planner and a forward planner.

The following runs the same example as the forward planner with and without the heuristic defined for the forward planner:
def test_regression_heuristic(thisproblem=problem1):
    print("\n***** REGRESSION NO HEURISTIC")
    print(SearcherMPP(Regression_STRIPS(thisproblem)).search())
    print("\n***** REGRESSION WITH HEURISTICS h1 and h2")
    print(SearcherMPP(Regression_STRIPS(thisproblem,maxh(h1,h2))).search())

if __name__ == "__main__":
    test_regression_heuristic()

Exercise 6.10 Try the regression planner with a heuristic function of just \( h_1 \) and with just \( h_2 \) (defined in Section 6.2.1). Explain how each one prunes or doesn’t prune the search space.

Exercise 6.11 Create a better heuristic than \( \text{heuristic fun} \) defined in Section 6.2.1

6.4 Planning as a CSP

To run the demo, in folder "aipython", load "stripsCSPPlanner.py", and copy and paste the commented-out example queries at the bottom of that file. This assumes Python 3.

Here we implement the CSP planner assuming there is a single action at each step. This creates a CSP that can use any of the CSP algorithms to solve (e.g., stochastic local search or arc consistency with domain splitting).

This assumes the same action representation as before; we do not consider factored actions (action features), nor do we implement state constraints.

```python
from cspProblem import Variable, CSP, Constraint

class CSP_from_STRIPS(CSP):
    """A CSP where:
    * CSP variables are constructed for each feature and time, and each
      action and time
    * the dynamics are specified by the STRIPS representation of actions
    """

def __init__(self, planning_problem, number_stages=2):
    prob_domain = planning_problem.prob_domain
    initial_state = planning_problem.initial_state
    goal = planning_problem.goal
    # self.action_vars[t] is the action variable for time t
    self.action_vars = [Variable(f"Action{t}" , prob_domain.actions)
                        for t in range(number_stages)]
    # feat_time_var[f][t] is the variable for feature f at time t
    feat_time_var = {feat: [Variable(f"{feat}_{t}" , dom)
                            for t in range(number_stages+1)]
                      for feat in planning_problem.init_features}
```
for (feat, dom) in
    prob_domain.feature_domain_dict.items()}

# initial state constraints:
constraints = [Constraint((feat_time_var[feat][0],), is_(val))
    for (feat, val) in initial_state.items()]

# goal constraints on the final state:
constraints += [Constraint((feat_time_var[feat][number_stages],),
    is_(val))
    for (feat, val) in goal.items()]

# precondition constraints:
constraints += [Constraint((feat_time_var[feat][t],
    self.action_vars[t]),
    if_(val, act)) # feat@t==val if action@t==act
    for act in prob_domain.actions
    for (feat, val) in act.preconds.items()
    for t in range(number_stages)]

# effect constraints:
constraints += [Constraint((feat_time_var[feat][t+1],
    self.action_vars[t]),
    if_(val, act)) # feat@t+1==val if action@t==act
    for act in prob_domain.actions
    for feat, val in act.effects.items()
    for t in range(number_stages)]

# frame constraints:
constraints += [Constraint((feat_time_var[feat][t],
    self.action_vars[t], feat_time_var[feat][t+1]),
    eq_if_not_in_((act for act in prob_domain.actions
    if feat in act.effects))
    for feat in prob_domain.feature_domain_dict
    for t in range(number_stages)]
variables = set(self.action_vars) | {feat_time_var[feat][t]
    for feat in prob_domain.feature_domain_dict
    for t in range(number_stages+1)}
CSP.__init__(self, variables, constraints)

def extract_plan(self, soln):
    return [soln[a] for a in self.action_vars]

The following methods return methods which can be applied to the particular
environment.

For example, is_(3) returns a function that when applied to 3, returns True
and when applied to any other value returns False. So is_(3)(3) returns True

http://aipython.org
Version 0.9.4
March 2, 2022
and \texttt{is\_\_} returns \texttt{False}.

Note that the underscore (\texttt{\_\_}) is part of the name; here we use it as the
convention that it is a function that returns a function. This uses two different
styles to define \texttt{is\_} and \texttt{if\_}; returning a function defined by \texttt{lambda} is equivalent
to returning the embedded function, except that the embedded function has a
name. The embedded function can also be given a docstring.

```python
stripsCSPPlanner.py — (continued)

def is\_\_(val):
    """returns a function that is true when it is applied to val."
    """
    #return lambda x: x == val
def is\_fun(x):
    return x == val
is\_fun.__name__ = "value_is_"+str(val)
return is\_fun

def if\_(v1,v2):
    """if the second argument is v2, the first argument must be v1""
    #return lambda x1,x2: x1==v1 if x2==v2 else True
def if\_fun(x1,x2):
    return x1==v1 if x2==v2 else True
if\_fun.__name__ = "if \_x2 is \_v2 then \_x1 is \_v1"
return if\_fun

def eq\_if\_not\_in\_(actset):
    """first and third arguments are equal if action is not in actset"
    # return lambda x1, a, x2: x1==x2 if a not in actset else True
def eq\_if\_not\_fun(x1, a, x2):
    return x1==x2 if a not in actset else True
eq\_if\_not\_fun.__name__ = "first and third arguments are equal if action
is not in \"\"+str(actset)
return eq\_if\_not\_fun
```

Putting it together, this returns a list of actions that solves the problem \texttt{prob}
for a given horizon. If you want to do more than just return the list of actions,
you might want to get it to return the solution. Or even enumerate the solutions
(by using \texttt{Search\_with\_AC\_from\_CSP}).

```python
stripsCSPPlanner.py — (continued)

def con\_plan(prob,horizon):
    """finds a plan for problem \texttt{prob} given horizon."
    """
    csp = CSP\_from\_STRIPS(prob, horizon)
sol = Con\_solver(csp).solve\_one()
return csp.extract\_plan(sol) if sol else sol
```

The following are some example queries.

```python
from searchGeneric import Searcher

http://aipython.org
```

Version 0.9.4  March 2, 2022
6.5 Partial-Order Planning

To run the demo, in folder "aipython", load "stripsPOP.py", and copy and paste the commented-out example queries at the bottom of that file.

http://aipython.org
A partial order planner maintains a partial order of action instances. An action instance consists of a name and an index. We need action instances because the same action could be carried out at different times.

```
from searchProblem import Arc, Search_problem
import random

class Action_instance(object):
    next_index = 0
    def __init__(self, action, index=None):
        if index is None:
            index = Action_instance.next_index
            Action_instance.next_index += 1
        self.action = action
        self.index = index

    def __str__(self):
        return str(self.action) + '#' + str(self.index)

    __repr__ = __str__ # __repr__ function is the same as the __str__ function
```

A node (as in the abstraction of search space) in a partial-order planner consists of:

- **actions**: a set of action instances.

- **constraints**: a set of \((a_1, a_2)\) pairs, where \(a_1\) and \(a_2\) are action instances, which represents that \(a_1\) must come before \(a_2\) in the partial order. There are a number of ways that this could be represented. Here we represent the set of pairs that are in transitive closure of the before relation. This lets us quickly determine whether some before relation is consistent with the current constraints.

- **agenda**: a list of \((s, a)\) pairs, where \(s\) is a \((var, val)\) pair and \(a\) is an action instance. This means that variable \(var\) must have value \(val\) before \(a\) can occur.

- **causal links**: a set of \((a_0, g, a_1)\) triples, where \(a_1\) and \(a_2\) are action instances and \(g\) is a \((var, val)\) pair. This holds when action \(a_0\) makes \(g\) true for action \(a_1\).
* actions is a set of action instances
* constraints a set of (a0,a1) pairs, representing a0<a1, closed under transitivity
* agenda list of (subgoal,action) pairs to be achieved, where subgoal is a (variable,value) pair
* causal_links is a set of (a0,g,a1) triples, where ai are action instances, and g is a (variable,value) pair

```python
self.actions = actions # a set of action instances
self.constraints = constraints # a set of (a0,a1) pairs
self.agenda = agenda # list of (subgoal,action) pairs to be achieved
self.causal_links = causal_links # set of (a0,g,a1) triples
```

```python
def __str__(self):
    return ('actions: ' + str({str(a) for a in self.actions}) +
    'constraints: ' + str({((str(a1),str(a2)) for (a1,a2) in self.constraints)} +
    '\nagenda: ' +
    str({((str(s),str(a)) for (s,a) in self.agenda}) +
    '\ncausal_links:' +
    str({((str(a0),str(g),str(a2)) for (a0,g,a2) in self.causal_links)})
```

`extract_plan` constructs a total order of action instances that is consistent with the partial order.

```python
def extract_plan(self):
    """returns a total ordering of the action instances consistent
    with the constraints.
    raises IndexError if there is no choice.
    """
    sorted_acts = []
    other_acts = set(self.actions)
    while other_acts:
        a = random.choice([a for a in other_acts if
                        all(((a1,a) not in self.constraints) for a1 in
                        other_acts)])
        sorted_acts.append(a)
        other_acts.remove(a)
    return sorted_acts
```

`POP_search_from_STRIPS` is an instance of a search problem. As such, we need to define the start nodes, the goal, and the neighbors of a node.

```python
from display import Displayable
class POP_search_from_STRIPS(Search_problem, Displayable):
    def __init__(self,planning_problem):
```

[http://aipython.org][1] Version 0.9.4 March 2, 2022
The `neighbors` method is a coroutine that enumerates the neighbors of a given node.
6.5. Partial-Order Planning

```python
new_cls = node.causal_links + [new_clink]
for consts3 in self.protect_all_cls(node.causal_links, new_a, consts2):
    for consts4 in self.protect_cl_for_actions(node.actions, consts3, new_clink):
        yield Arc(node,
                  POP_node(new_actions, consts4, new_agenda1, new_cls),
                  cost=1)
```

Given a causal link \((a_0, \text{subgoal}, a_1)\), the following method protects the causal link from each action in \(\text{actions}\). Whenever an action deletes \(\text{subgoal}\), the action needs to be before \(a_0\) or after \(a_1\). This method enumerates all constraints that result from protecting the causal link from all actions.

```python
def protect_cl_for_actions(self, actions, constrs, clink):
    """yields constraints that extend constrs and
    protect causal link \((a_0, \text{subgoal}, a_1)\)
    for each action in actions
    ""
    if actions:
        a = actions[0]
        rem_actions = actions[1:]
        a0, subgoal, a1 = clink
        if a != a0 and a != a1 and self.deletes(a, subgoal):
            if self.possible((a, a0), constrs):
                new_const = self.add_constraint((a, a0), constrs)
                for e in self.protect_cl_for_actions(rem_actions, new_const, clink):
                    yield e  # could be "yield from"
            if self.possible((a1, a), constrs):
                new_const = self.add_constraint((a1, a), constrs)
                for e in self.protect_cl_for_actions(rem_actions, new_const, clink):
                    yield e
        else:
            for e in self.protect_cl_for_actions(rem_actions, constrs, clink):
                yield e
    else:
        yield constrs
```

Given an action \(\text{act}\), the following method protects all the causal links in \(\text{clinks}\) from \(\text{act}\). Whenever \(\text{act}\) deletes \(\text{subgoal}\) from some causal link \((a_0, \text{subgoal}, a_1)\), the action \(\text{act}\) needs to be before \(a_0\) or after \(a_1\). This method enumerates all constraints that result from protecting the causal links from \(\text{act}\).

```python
def protect_all_cls(self, clinks, act, constrs):
    """yields constraints that protect all causal links from \(\text{act}\)"
    if clinks:
```

http://aipython.org
The following methods check whether an action (or action instance) achieves or deletes some subgoal.

```python
def achieves(self, action, subgoal):
    var, val = subgoal
    return var in self.effects(action) and self.effects(action)[var] == val

def deletes(self, action, subgoal):
    var, val = subgoal
    return var in self.effects(action) and self.effects(action)[var] != val

def effects(self, action):
    """returns the variable:value dictionary of the effects of action. works for both actions and action instances""
    if isinstance(action, Action_instance):
        action = action.action
    if action == "start":
        return self.planning_problem.initial_state
    elif action == "finish":
        return {}
    else:
        return action.effects
```

The constraints are represented as a set of pairs closed under transitivity. Thus if (a, b) and (b, c) are the list, then (a, c) must also be in the list. This means that adding a new constraint means adding the implied pairs, but querying whether some order is consistent is quick.
6.5. Partial-Order Planning

```python
    return const
todo = [pair]
newconst = const.copy()
while todo:
x0, x1 = todo.pop()
newconst.add((x0, x1))
    for x, y in newconst:
        if x == x1 and (x0, y) not in newconst:
todo.append((x0, y))
        if y == x0 and (x, x1) not in newconst:
todo.append((x, x1))
    return newconst

def possible(self, pair, constraint):
    (x, y) = pair
    return (y, x) not in constraint
```

Some code for testing:

```python
from searchBranchAndBound import DF_branch_and_bound
from searchMPP import SearcherMPP
from stripsProblem import problem0, problem1, problem2, blocks1, blocks2, blocks3

rplanning0 = POP_search_from_STRIPS(problem0)
rplanning1 = POP_search_from_STRIPS(problem1)
rplanning2 = POP_search_from_STRIPS(problem2)
searcher0 = DF_branch_and_bound(rplanning0, 5)
searcher0a = SearcherMPP(rplanning0)
searcher1 = DF_branch_and_bound(rplanning1, 10)
searcher1a = SearcherMPP(rplanning1)
searcher2 = DF_branch_and_bound(rplanning2, 10)
searcher2a = SearcherMPP(rplanning2)

# Try one of the following searchers
# a = searcher0.search()
# a = searcher0a.search()
# a = searcher1.search()
# a = searcher1a.search()
# a = searcher2.search()
# a = searcher2a.search()
```

http://aipython.org

Version 0.9.4
March 2, 2022
Supervised Machine Learning

This chapter is the first on machine learning. It covers the following topics:

- Data: how to load it, training and test sets
- Features: many of the features come directly from the data. Sometimes it is useful to construct features, e.g. \( \text{height} > 1.9 \text{m} \) might be a Boolean feature constructed from the real-values feature \( \text{height} \). The next chapter is about neural networks and how to learn features; in this chapter we construct explicitly in what is often known a feature engineering.
- Learning with no input features: this is the base case of many methods. What should we predict if we have no input features? This provides the base cases for many algorithms (e.g., decision tree algorithm) and baselines that more sophisticated algorithms need to beat. It also provides ways to test various predictors.
- Decision tree learning: one of the classic and simplest learning algorithms, which is the basis of many other algorithms.
- Cross validation and parameter tuning: methods to prevent overfitting.
- Linear regression and classification: other classic and simple techniques that often work well (particularly combined with feature learning or engineering).
- Boosting: combining simpler learning methods to make even better learners.

A good source of classic datasets is the UCI Machine Learning Repository [Lichman, 2013][?]. The SPECT and car datasets are from this repository.
## 7. Supervised Machine Learning

<table>
<thead>
<tr>
<th>Dataset</th>
<th># Examples</th>
<th>#Columns</th>
<th>Input Types</th>
<th>Target Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECT</td>
<td>267</td>
<td>23</td>
<td>Boolean</td>
<td>Boolean</td>
</tr>
<tr>
<td>IRIS</td>
<td>150</td>
<td>5</td>
<td>real</td>
<td>categorical</td>
</tr>
<tr>
<td>carbool</td>
<td>1728</td>
<td>7</td>
<td>categorical/real</td>
<td>real</td>
</tr>
<tr>
<td>holiday</td>
<td>32</td>
<td>6</td>
<td>Boolean</td>
<td>Boolean</td>
</tr>
<tr>
<td>mail_reading</td>
<td>28</td>
<td>5</td>
<td>Boolean</td>
<td>Boolean</td>
</tr>
<tr>
<td>simp_regr</td>
<td>7</td>
<td>2</td>
<td>numerical</td>
<td>numerical</td>
</tr>
</tbody>
</table>

Figure 7.1: Some of the datasets used here. MLR is UCI Machine Learning Repository.

### 7.1 Representations of Data and Predictions

The code uses the following definitions and conventions:

- A **data set** is an enumeration of examples.

- An **example** is a list (or tuple) of values. The values can be numbers or strings.

- A **feature** is a function from examples into the range of the feature. Each feature $f$ also has the following attributes:

  $f$.ftype the type of $f$; one of: "boolean", "categorical", "numeric"

  $f$.frange the range of $f$, represented as a list

  $f$.__doc__ the docstring, a string description of $f$ (for printing).

Thus for example, a **Boolean feature** is a function from the examples into \{False, True\}. So, if $f$ is a Boolean feature, $f.frange == \{False, True\}$, and if $e$ is an example, $f(e)$ is either True or False.

```python
import math, random, statistics
import csv
from display import Displayable
from utilities import argmax

boolean = [False, True]
```

When creating a data set, we partition the data into a training set ($train$) and a test set ($test$). The target feature is the feature that we are making a prediction of. A dataset $ds$ has the following attributes

- $ds.train$ a list of the training examples
- $ds.test$ a list of the test examples
- $ds.target_index$ the index of the target

[http://aipython.org](http://aipython.org)  
Version 0.9.4  
March 2, 2022
ds.target the feature corresponding to the target (a function as described above)

ds.input_features a list of the input features

```python
class Data_set(Displayable):
    """ A data set consists of a list of training data and a list of test data. """

    def __init__(self, train, test=None, prob_test=0.20, target_index=0,
                 header=None, target_type=None, seed=None): #12345):
        """A dataset for learning.
        train is a list of tuples representing the training examples
        test is the list of tuples representing the test examples
        if test is None, a test set is created by selecting each
        example with probability prob_test
        target_index is the index of the target.
        If negative, it counts from right.
        If target_index is larger than the number of properties,
        there is no target (for unsupervised learning)
        header is a list of names for the features
        target_type is either None for automatic detection of target type
        or one of "numerical", "boolean", "cartegorical"
        seed is for random number; None gives a different test set each time
        """
        if seed: # given seed makes partition consistent from run-to-run
            random.seed(seed)
        if test is None:
            train, test = partition_data(train, prob_test)
        self.train = train
        self.test = test

        self.display(1,"Training set has",len(train),"examples. Number of columns: ",
                     "\n                     \{\n                     \len(e) \for \ e \in \ train\n                     \}
        self.display(1,"Test set has",len(test),"examples. Number of columns: ",
                     \n                     \{\n                     \len(e) \for \ e \in \ test\n                     \}
        self.prob_test = prob_test
        self.num_properties = len(self.train[0])
        if target_index < 0: #allows for -1, -2, etc.
            self.target_index = self.num_properties + target_index
        else:
            self.target_index = target_index
        self.header = header
        self.domains = [set() for i in range(self.num_properties)]
        for example in self.train:
            for ind, val in enumerate(example):
                self.domains[ind].add(val)
        self.conditions_cache = {} # cache for computed conditions
```

[http://aipython.org](http://aipython.org) Version 0.9.4 March 2, 2022
7. Supervised Machine Learning

```python
self.create_features()
if target_type:
    self.target.ftype = target_type
self.display(1, "There are", len(self.input_features), "input features")

def __str__(self):
    if self.train and len(self.train) > 0:
        return ("Data: "+
            str(len(self.train))+" training examples, " +
            str(len(self.test))+" test examples, "+
            str(len(self.train[0]))+" features.")
    else:
        return ("Data: "+
            str(len(self.train))+" training examples, " +
            str(len(self.test))+" test examples.")
```

A **feature** is a function that takes an example and returns a value in the range of the feature. Each feature has a **frange**, which gives the range of the feature, and an **ftype** that gives the type, one of “boolean”, “numeric” or “categorical”.

```python
def create_features(self):
    """create the set of features
    ""
    self.target = None
    self.input_features = []
    for i in range(self.num_properties):
        def feat(e,index=i):
            return e[index]
        if self.header:
            feat.__doc__ = self.header[i]
        else:
            feat.__doc__ = "e["+str(i)+"]"
        feat.frange = list(self.domains[i])
        feat.ftype = self.infer_type(feat.frange)
        if i == self.target_index:
            self.target = feat
        else:
            self.input_features.append(feat)
```

We try to infer the type of each feature. Sometimes this can be wrong, (e.g., when the numbers are really categorical) and so needs to be set explicitly.

```python
def infer_type(self,domain):
    """Infers the type of a feature with domain
    ""
    if all(v in {True,False} for v in domain):
        return "boolean"
    if all(isinstance(v,(float,int)) for v in domain):
        return "numeric"
```

http://aipython.org  Version 0.9.4  March 2, 2022
7.1. Representations of Data and Predictions

7.1.1 Creating Boolean Conditions from Features

Some of the algorithms require Boolean input features or features with range \{0,1\}. In order to be able to use these algorithms on datasets that allow for arbitrary domains of input variables, we construct Boolean conditions from the attributes.

There are 3 cases:

- When the range only has two values, we designate one to be the “true” value.

- When the values are all numeric, we assume they are ordered (as opposed to just being some classes that happen to be labelled with numbers) and construct Boolean features for splits of the data. That is, the feature is \( e[\text{ind}] < \text{cut} \) for some value \text{cut}. We choose a number of \text{cut} values, up to a maximum number of cuts, given by \text{max_num_cuts}.

- When the values are not all numeric, we create an indicator function for each value. An indicator function for a value returns true when that value is given and false otherwise. Note that we can’t create an indicator function for values that appear in the test set but not in the training set because we haven’t seen the test set. For the examples in the test set with a value that doesn’t appear in the training set for that feature, the indicator functions all return false.

There is also an option to only create Boolean features from categorical input features.

```
def conditions(self, max_num_cuts=8, categorical_only = False):
    """returns a set of boolean conditions from the input features
    max_num_cuts is the maximum number of cute for numerical features
    categorical_only is true if only categorical features are made binary
    """
    if (max_num_cuts, categorical_only) in self.conditions_cache:
        return self.conditions_cache[(max_num_cuts, categorical_only)]
    conds = []
    for ind,frange in enumerate(self.domains):
        if ind != self.target_index and len(frange)>1:
            if len(frange) == 2:
                # two values, the feature is equality to one of them.
                true_val = list(frange)[1]  # choose one as true
                def feat(e, i=ind, tv=true_val):
                    return e[i]==tv
```
if self.header:
    feat.__doc__ = f"{self.header[ind]}=={true_val}"
else:
    feat.__doc__ = f"e[{ind}]=={true_val}"
feat.frange = boolean
feat.ftype = "boolean"
conds.append(feat)

elif all(isinstance(val,(int, float))) for val in frange):
    if categorical_only: # numerical, don't make cuts
        def feat(e, i=ind):
            return e[i]
        feat.__doc__ = f"e[{ind}]"
        conds.append(feat)
    else:
        # all numeric, create cuts of the data
        sorted_frange = sorted(frange)
        num_cuts = min(max_num_cuts, len(frange))
        cut_positions = [len(frange)*i//num_cuts for i in range(1,num_cuts)]
        for cut in cut_positions:
            cutat = sorted_frange[cut]
            def feat(e, ind_=ind, cutat=cutat):
                return e[ind_] < cutat
            if self.header:
                feat.__doc__ = self.header[ind] + "<" + str(cutat)
            else:
                feat.__doc__ = f"e[{ind}]<=\{cutat}\n            feat.frange = boolean
            feat.ftype = "boolean"
            conds.append(feat)
        else:
            # create an indicator function for every value
            for val in frange:
                def feat(e, ind_=ind, val_=val):
                    return e[ind_] == val_
                if self.header:
                    feat.__doc__ = self.header[ind] + "==\{val\}
                else:
                    feat.__doc__ = f"e[{ind}]==\{val}\n                feat.frange = boolean
                feat.ftype = "boolean"
                conds.append(feat)
        self.conditions_cache[(max_num_cuts, categorical_only)] = conds
        return conds

Exercise 7.1 Change the code so that it splits using $e[\text{ind}] \leq \text{cut}$ instead of $e[\text{ind}] < \text{cut}$. Check boundary cases, such as 3 elements with 2 cuts. As a test case, make sure that when the range is the 30 integers from 100 to 129, and you want 2 cuts, the resulting Boolean features should be $e[\text{ind}] \leq 109$ and $e[\text{ind}] \leq 119$ to make
7.1. Representations of Data and Predictions

Exercise 7.2 This splits on whether the feature is less than one of the values in the training set. Sam suggested it might be better to split between the values in the training set, and suggested using

$$cutat = \frac{\text{sorted\_frange}[\text{cut}] + \text{sorted\_frange}[\text{cut} - 1]}{2}$$

Why might Sam have suggested this? Does this work better? (Try it on a few data sets).

7.1.2 Evaluating Predictions

A predictor is a function that takes an example and makes a prediction on the values of the target features.

A loss takes a prediction and the actual value and returns a non-negative real number; lower is better. The error for a dataset is either the mean loss, or sometimes the sum of the losses. When reporting results the mean is usually used. When it is the sum, this will be made explicit.

The function evaluate_dataset returns the average error for each example, where the error for each example depends on the evaluation criteria. Here we consider three evaluation criteria, the squared error (average of the square of the difference between the actual and predicted values), absolute errors (average of the absolute difference between the actual and predicted values) and the log loss (the average negative log-likelihood, which can be interpreted as the number of bits to describe an example using a code based on the prediction treated as a probability).

```
def evaluate_dataset(self, data, predictor, error_measure):
    """Evaluates predictor on data according to the error_measure
    predictor is a function that takes an example and returns a
    prediction for the target features.
    error_measure(prediction,actual) \to non-negative real
    """
    if data:
        try:
            value = statistics.mean(error_measure(predictor(e),
                        self.target(e))
            for e in data)
        except ValueError: # if error_measure gives an error
            return float("inf") # infinity
        return value
    else:
        return math.nan # not a number
```

The following evaluation criteria are defined. This is defined using a class, Evaluate but no instances will be created. Just use Evaluate.squared_loss etc. (Please keep the \_doc\_ strings a consistent length as they are used in tables.)

http://aipython.org  Version 0.9.4  March 2, 2022
The prediction is either a real value or a \{value : probability\} dictionary or a list. The actual is either a real number or a key of the prediction.

```python
class Evaluate(object):
    """A container for the evaluation measures"""
    def squared_loss(prediction, actual):
        """squared loss"
        if isinstance(prediction, (list, dict)):
            return (1-prediction[actual])**2 # the correct value is 1
        else:
            return (prediction-actual)**2
    def absolute_loss(prediction, actual):
        """absolute loss"
        if isinstance(prediction, (list, dict)):
            return abs(1-prediction[actual]) # the correct value is 1
        else:
            return abs(prediction-actual)
    def log_loss(prediction, actual):
        """log loss (bits)"
        try:
            if isinstance(prediction, (list, dict)):
                return -math.log2(prediction[actual])
            else:
                return -math.log2(prediction) if actual==1 else -math.log2(1-prediction)
        except ValueError:
            return float("inf") # infinity
    def accuracy(prediction, actual):
        """accuracy"
        if isinstance(prediction, dict):
            prev_val = prediction[actual]
            return 1 if all(prev_val >= v for v in prediction.values()) else 0
        if isinstance(prediction, list):
            prev_val = prediction[actual]
            return 1 if all(prev_val >= v for v in prediction) else 0
        else:
            return 1 if abs(actual-prediction) <= 0.5 else 0
all_criteria = [accuracy, absolute_loss, squared_loss, log_loss]
```

### 7.1.3 Creating Test and Training Sets

The following method partitions the data into a training set and a test set. Note that this does not guarantee that the test set will contain exactly a proportion

[http://aipython.org](http://aipython.org)
of the data equal to \( \text{prob\_test} \).

[An alternative is to use \texttt{random.sample()} which can guarantee that the test set will contain exactly a particular proportion of the data. However this would require knowing how many elements are in the data set, which we may not know, as \texttt{data} may just be a generator of the data (e.g., when reading the data from a file).]

```python
    def partition_data(data, prob_test=0.30):
        """partitions the data into a training set and a test set, where prob_test is the probability of each example being in the test set. """
        train = []
        test = []
        for example in data:
            if random.random() < prob_test:
                test.append(example)
            else:
                train.append(example)
        return train, test
```

### 7.1.4 Importing Data From File

A data set is typically loaded from a file. The default here is that it loaded from a CSV (comma separated values) file, although the separator can be changed. This assumes that all lines that contain the separator are valid data (so we only include those data items that contain more than one element). This allows for blank lines and comment lines that do not contain the separator. However, it means that this method is not suitable for cases where there is only one feature.

Note that \texttt{data\_all} and \texttt{data\_tuples} are generators. \texttt{data\_all} is a generator of a list of list of strings. This version assumes that CSV files are simple. The standard \texttt{csv} package, that allows quoted arguments, can be used by uncommenting the line for \texttt{data\_all} and commenting out the following line. \texttt{data\_tuples} contains only those lines that contain the delimiter (others lines are assumed to be empty or comments), and tries to convert the elements to numbers whenever possible.

This allows for some of the columns to be included; specified by \texttt{include\_only}. Note that if \texttt{include\_only} is specified, the target index is the index for the included columns, not the original columns.

```python
    class Data_from_file(Data_set):
        def __init__(self, file_name, separator=',', num_train=None, prob_test=0.3, has_header=False, target_index=0, boolean_features=True, categorical=[], target_type=None, include_only=None, seed=None): #seed=12345):
```

http://aipython.org
create a dataset from a file

separater is the character that separates the attributes
num_train is a number specifying the first num_train tuples are training, or None
prob_test is the probability an example should in the test set (if num_train is None)
has_header is True if the first line of file is a header
target_index specifies which feature is the target
boolean_features specifies whether we want to create Boolean features
   (if False, it uses the original features).
categorical is a set (or list) of features that should be treated as categorical
target_type is either None for automatic detection of target type
   or one of "numerical", "boolean", "cartegorical"
include_only is a list or set of indexes of columns to include

self.boolean_features = boolean_features
with open(file_name,'r',newline='') as csvfile:
  self.display(1,"Loading",file_name)
  # data_all = csv.reader(csvfile,delimiter=separator) # for more complicated CSV files
  data_all = (line.strip().split(separator) for line in csvfile)
  if include_only is not None:
    data_all = ([v for (i,v) in enumerate(line) if i in include_only]
                for line in data_all)
  if has_header:
    header = next(data_all)
  else:
    header = None
  data_tuples = (interpret_elements(d) for d in data_all if len(d)>1)
  if num_train is not None:
    # training set is divided into training then text examples
    # the file is only read once, and the data is placed in appropriate list
    train = []
    for i in range(num_train): # will give an error if insufficient examples
      train.append(next(data_tuples))
    test = list(data_tuples)
  else: # randomly assign training and test examples
    Data_set.__init__(self,train, test=test,
      target_index=target_index,header=header)
else:
  Data_set.__init__(self,data_tuples, test=None,
                      prob_test=prob_test,
                      target_index=target_index, header=header,
                      seed=seed, target_type=target_type)

The following class is used for datasets where the training and test are in dif-
7.1. Representations of Data and Predictions

ferent files

```python
class Data_from_files(Data_set):
    def __init__(self, train_file_name, test_file_name, separator='\',
                 has_header=False, target_index=\0, boolean_features=True,
                 categorical=[], target_type= None, include_only=None):
        r"""create a dataset from separate training and file
        separator is the character that separates the attributes
        num_train is a number specifying the first num_train tuples are
        training, or None
        prob_test is the probability an example should in the test set (if
        num_train is None)
        has_header is True if the first line of file is a header
        target_index specifies which feature is the target
        boolean_features specifies whether we want to create Boolean
        features
        (if False, it uses the original features).
        categorical is a set (or list) of features that should be treated
        as categorical
        target_type is either None for automatic detection of target type
        or one of "numerical", "boolean", "categorical"
        include_only is a list or set of indexes of columns to include
        """
        self.boolean_features = boolean_features
        with open(train_file_name,'r',newline='') as train_file:
            with open(test_file_name,'r',newline='') as test_file:
                # data_all = csv.reader(csvfile,delimiter=separator) # for more
                # complicated CSV files
                train_data = (line.strip().split(separator) for line in
                              train_file)
                test_data = (line.strip().split(separator) for line in
                             test_file)
                if include_only is not None:
                    train_data = ([v for (i,v) in enumerate(line) if i in
                                   include_only]
                                   for line in train_data)
                    test_data = ([v for (i,v) in enumerate(line) if i in
                                   include_only]
                                   for line in test_data)
                if has_header: # this assumes the training file has a header
                    header = next(train_data)
                else:
                    header = None
                train_tuples = [interpret_elements(d) for d in train_data if
                                 len(d)>1]
                test_tuples = [interpret_elements(d) for d in test_data if
                               len(d)>1]
                Data_set.__init__(self,train_tuples, test_tuples,
                                   target_index=target_index, header=header)
```

http://aipython.org
When reading from a file all of the values are strings. This next method tries to convert each value into a number (an int or a float) or Boolean, if it is possible.

```
def interpret_elements(str_list):
    """make the elements of string list str_list numerical if possible.
    Otherwise remove initial and trailing spaces.
    """
    res = []
    for e in str_list:
        try:
            res.append(int(e))
        except ValueError:
            try:
                res.append(float(e))
            except ValueError:
                se = e.strip()
                if se in ["True", "true", "TRUE"]: res.append(True)
                if se in ["False", "false", "FALSE"]: res.append(False)
                else: res.append(e.strip())
    return res
```

7.1.5 Augmented Features

Sometimes we want to augment the features with new features computed from the old features (e.g. the product of features). Here we allow the creation of a new dataset from an old dataset but with new features. Note that special cases of these are kernels; mapping the original feature space into a new space, which allow a neat way to do learning in the augmented space (the “kernel trick”). This is beyond the scope of AIPython; those interested should read about support vector machines.

A feature is a function of examples. A unary feature constructor takes a feature and returns a new feature. A binary feature combiner takes two features and returns a new feature.

```
class Data_set_augmented(Data_set):
    def __init__(self, dataset, unary_functions=[], binary_functions=[], include_orig=True):
        """creates a dataset like dataset but with new features
unary_function is a list of unary feature constructors
binary_function is a list of binary feature combiners.
include_orig specifies whether the original features should be included
```

http://aipython.org
337

self.orig_dataset = dataset
self.unary_functions = unary_functions
self.binary_functions = binary_functions
self.include_orig = include_orig
self.target = dataset.target
Data_set.__init__(self,dataset.train, test=dataset.test,
                   target_index = dataset.target_index)

346
def create_features(self):
    if self.include_orig:
        self.input_features = self.orig_dataset.input_features.copy()
    else:
        self.input_features = []
    for u in self.unary_functions:
        for f in self.orig_dataset.input_features:
            self.input_features.append(u(f))
    for b in self.binary_functions:
        for f1 in self.orig_dataset.input_features:
            for f2 in self.orig_dataset.input_features:
                if f1 != f2:
                    self.input_features.append(b(f1,f2))

The following are useful unary feature constructors and binary feature combiner.

```
def square(f):
    """a unary feature constructor to construct the square of a feature
    """
    def sq(e):
        return f(e)**2
    sq.__doc__ = f.__doc__ + "**2"
    return sq

def power_feat(n):
    """given n returns a unary feature constructor to construct the nth power of a feature.
    e.g., power_feat(2) is the same as square, defined above
    """
    def fn(f,n=n):
        def pow(e,n=n):
            return f(e)**n
        pow.__doc__ = f.__doc__ + "**" + str(n)
        return pow
    return fn

def prod_feat(f1,f2):
    """a new feature that is the product of features f1 and f2
    """
    def feat(e):
```
Supervised Machine Learning

```python
    return f1(e)*f2(e)
    feat.__doc__ = f1.__doc__+"*"+f2.__doc__
    return feat

def eq_feat(f1,f2):
    """a new feature that is 1 if f1 and f2 give same value
    """
    def feat(e):
        return 1 if f1(e)==f2(e) else 0
    feat.__doc__ = f1.__doc__+"=="+f2.__doc__
    return feat

def neq_feat(f1,f2):
    """a new feature that is 1 if f1 and f2 give different values
    """
    def feat(e):
        return 1 if f1(e)!=f2(e) else 0
    feat.__doc__ = f1.__doc__+"!="+f2.__doc__
    return feat
```

Example:

```python
# from learnProblem import Data_set_augmented, prod_feat
# data = Data_from_file('data/holiday.csv', num_train=19, target_index=-1)
# data = Data_from_file('data/iris.data', prob_test=1/3, target_index=-1)
# Data = Data_from_file('data/SPECT.csv', prob_test=0.5, target_index=0)
# dataplus = Data_set_augmented(data,[],[prod_feat])
# dataplus = Data_set_augmented(data,[],[prod_feat,neq_feat])
```

Exercise 7.3 For symmetric properties, such as product, we don’t need both \( f_1 \times f_2 \) as well as \( f_2 \times f_1 \) as extra properties. Allow the user to be able to declare feature constructors as symmetric (by associating a Boolean feature with them). Change `construct_features` so that it does not create both versions for symmetric combiners.

7.2 Generic Learner Interface

A learner takes a dataset (and possibly other arguments specific to the method). To get it to learn, we call the `learn()` method. This implements `Displayable` so that we can display traces at multiple levels of detail (and perhaps with a GUI).
7.3. Learning With No Input Features

If we make the same prediction for each example, what prediction should we make? This can be used as a naive baseline; if a more sophisticated method does not do better than this, it is not useful. This also provides the base case for some methods, such as decision-tree learning.

To run demo to compare different prediction methods on various evaluation criteria, in folder "aipython", load "learnNoInputs.py", using e.g., ipython -i learnNoInputs.py, and it prints some test results.

There are a few alternatives as to what could be allowed in a prediction:

- a point prediction, where we are only allowed to predict one of the values of the feature. For example, if the values of the feature are \{0, 1\} we are only allowed to predict 0 or 1 or of the values are ratings in \{1, 2, 3, 4, 5\}, we can only predict one of these integers.

- a point prediction, where we are allowed to predict any value. For example, if the values of the feature are \{0, 1\} we may be allowed to predict 0.3, 1, or even 1.7. For all of the criteria we can imagine, there is no point in predicting a value greater than 1 or less that zero (but that doesn’t mean we can’t), but it is often useful to predict a value between 0 and 1. If the values are ratings in \{1, 2, 3, 4, 5\}, we may want to predict 3.4.

- a probability distribution over the values of the feature. For each value \(v\), we predict a non-negative number \(p_v\), such that the sum over all predictions is 1.

For regression, we do the first of these. For classification, we do the second. The third can be implemented by having multiple indicator functions for the target.

Here are some prediction functions that take in an enumeration of values, a domain, and returns a value or dictionary of \{value : prediction\}. Note that \(\text{cmedian}\) returns one of middle values when there are an even number of examples, whereas \(\text{median}\) gives the average of them (and so \(\text{cmedian}\) is applicable for ordinals that cannot be considered cardinal values). Similarly, \(\text{cmode}\) picks one of the values when more than one value has the maximum number of elements.
7. Supervised Machine Learning

```python
from learnProblem import Evaluate
import math, random, collections, statistics
import utilities # argmax for (element,value) pairs

class Predict(object):
    """The class of prediction methods for a list of values.
    Please make the doc strings the same length, because they are used in
tables.
    Note that we don't need self argument, as we are creating Predict
    objects,
    To use call Predict.laplace(data) etc.""

    ### The following return a distribution over values (for classification)
def empirical(data, domain=[0,1], icount=0):
        "empirical dist "
        # returns a distribution over values
        counts = {v:icount for v in domain}
        for e in data:
            counts[e] += 1
        s = sum(counts.values())
        return {k:v/s for (k,v) in counts.items()}

def bounded_empirical(data, domain=[0,1], bound=0.01):
        "bounded empirical"
        return {k:min(max(v,bound),1-bound) for (k,v) in
                Predict.empirical(data, domain).items()}

def laplace(data, domain=[0,1]):
        "Laplace " # for categorical data
        return Predict.empirical(data, domain, icount=1)

def cmode(data, domain=[0,1]):
        "mode " # for categorical data
        md = statistics.mode(data)
        return {v: 1 if v==md else 0 for v in domain}

def cmedian(data, domain=[0,1]):
        "median " # for categorical data
        md = statistics.median_low(data) # always return one of the values
        return {v: 1 if v==md else 0 for v in domain}

    ### The following return a single prediction (for regression). domain
    is ignored.
def mean(data, domain=[0,1]):
        "mean "
        # returns a real number
        return statistics.mean(data)

http://aipython.org  Version 0.9.4  March 2, 2022
7.3. Learning With No Input Features

```python
def rmean(data, domain=[0,1], mean0=0, pseudo_count=1):
    "regularized mean"
    # returns a real number.
    # mean0 is the mean to be used for 0 data points
    # With mean0=0.5, pseudo_count=2, same as laplace for [0,1] data
    # this works for enumerations as well as lists
    sum = mean0 * pseudo_count
    count = pseudo_count
    for e in data:
        sum += e
        count += 1
    return sum/count

def mode(data, domain=[0,1]):
    "mode"
    return statistics.mode(data)

def median(data, domain=[0,1]):
    "median"
    return statistics.median(data)

all = [empirical, mean, rmean, bounded_empirical, laplace, cmode, mode, median, cmedian]

# The following suggests appropriate predictions as a function of the
target type
select = {
    "boolean": [empirical, bounded_empirical, laplace, cmode, cmode, median, cmedian],
    "categorical": [empirical, bounded_empirical, laplace, cmode, cmedian],
    "numeric": [mean, rmean, mode, median]
}

7.3.1 Evaluation

To evaluate a point prediction, we first generate some data from a simple (Bernoulli) distribution, where there are two possible values, 0 and 1 for the target feature. Given prob, a number in the range [0,1], this generate some training and test data where prob is the probability of each example being 1. To generate a 1 with probability prob, we generate a random number in range [0,1] and return 1 if that number is less than prob. A prediction is computed by applying the predictor to the training data, which is evaluated on the test set. This is repeated num_samples times.

Let’s evaluate the predictions of the possible selections according to the different evaluation criteria, for various training sizes.
```

```
results = {predictor: {error_measure: 0 for error_measure in error_measures}
           for predictor in Predict.all}
for sample in range(num_samples):
    prob = random.random()
    training = [1 if random.random() < prob else 0 for i in range(train_size)]
    test = [1 if random.random() < prob else 0 for i in range(test_size)]
    for predictor in Predict.all:
        prediction = predictor(training)
        results[predictor][error_measure] += sum(error_measure(prediction, actual) for actual in test) / test_size
print(f"For training size {train_size}:")
print(" Predictor\t".join(error_measure.__doc__ for error_measure in error_measures), sep="\t")
for predictor in Predict.all:
    print(f" {predictor.__doc__}",
          "\t".join("{:.7f}".format(results[predictor][error_measure] / num_samples)
                      for error_measure in error_measures), sep="\t")

if __name__ == "__main__":
    test_no_inputs()
7.4 Decision Tree Learning

To run the decision tree learning demo, in folder "aipython", load "learnDT.py", using e.g., ipython -i learnDT.py, and it prints some test results. To try more examples, copy and paste the commented-out commands at the bottom of that file. This requires Python 3 with matplotlib.

The decision tree algorithm does binary splits, and assumes that all input features are binary functions of the examples. It stops splitting if there are no input features, the number of examples is less than a specified number of examples or all of the examples agree on the target feature.

```python
from learnProblem import Learner, Evaluate
from learnNoInputs import Predict
import math

class DT_learner(Learner):
    def __init__(self,
        dataset,
        split_to_optimize=Evaluate.log_loss, # to minimize for at each split
        leaf_prediction=Predict.empirical, # what to use for value at leaves
        train=None, # used for cross validation
        max_num_cuts=8, # maximum number of conditions to split a numerical feature into
        gamma=1e-7 , # minimum improvement needed to expand a node
        min_child_weight=10):
        self.dataset = dataset
        self.target = dataset.target
        self.split_to_optimize = split_to_optimize
        self.leaf_prediction = leaf_prediction
        self.max_num_cuts = max_num_cuts
        self.gamma = gamma
        self.min_child_weight = min_child_weight

        if train is None:
            self.train = self.dataset.train
        else:
            self.train = train

    def learn(self, max_num_cuts=8):
        """learn a decision tree""
        return self.learn_tree(self.dataset.conditions(self.max_num_cuts),
                                self.train)
```

The main recursive algorithm, takes in a set of input features and a set of training data. It first decides whether to split. If it doesn’t split, it makes a point prediction, ignoring the input features.

http://aipython.org
It splits unless:

- there are no more input features
- there are fewer examples than \texttt{min\_number\_examples},
- all the examples agree on the value of the target, or
- the best split makes all examples in the same partition.

If it splits, it selects the best split according to the evaluation criterion (assuming that is the only split it gets to do), and returns the condition to split on (in the variable \texttt{split}) and the corresponding partition of the examples.

```python
def learn_tree(self, conditions, data_subset):
    """returns a decision tree
    conditions is a set of possible conditions
    data_subset is a subset of the data used to build this (sub)tree
    """
    self.display(2,f"learn_tree with {len(conditions)} features and
    {len(data_subset)} examples")
    split, partn = self.select_split(conditions, data_subset)
    if split is None: # no split; return a point prediction
        prediction = self.leaf_value(data_subset, self.target.frange)
        self.display(2,f"leaf prediction for {len(data_subset)} examples is {prediction}")
        def leaf_fun(e):
            return prediction
        leaf_fun.__doc__ = str(prediction)
        leaf_fun.num_leaves = 1
        return leaf_fun
    else: # a split succeeded
        false_examples, true_examples = partn
        rem_features = [fe for fe in conditions if fe != split]
        self.display(2,"Splitting on",split.__doc__,"with examples split",
        len(true_examples),":",len(false_examples))
        true_tree = self.learn_tree(rem_features,true_examples)
        false_tree = self.learn_tree(rem_features,false_examples)
        def fun(e):
            if split(e):
                return true_tree(e)
            else:
                return false_tree(e)
        fun.__doc__ = ("if "+split.__doc__+") then ("+true_tree.__doc__+
        ") else ("+false_tree.__doc__")"
```
fun.num_leaves = true_tree.num_leaves + false_tree.num_leaves

    return fun

---

def leaf_value(self, egs, domain):
    return self.leaf_prediction((self.target(e) for e in egs), domain)

def select_split(self, conditions, data_subset):
    """finds best feature to split on.

    conditions is a non-empty list of features.
    returns feature, partition
    where feature is an input feature with the smallest error as
    judged by split_to_optimize or
    feature==None if there are no splits that improve the error
    partition is a pair (false_examples, true_examples) if feature is
    not None
    """
    best_feat = None # best feature
    best_error = self.sum_losses(data_subset) - self.gamma
    self.display(3," no split has
    error",best_error,"with",len(conditions),"conditions")
    best_partition = None
    for feat in conditions:
        false_examples, true_examples = partition(data_subset,feat)
        if min(len(false_examples),len(true_examples))>=self.min_child_weight:
            err = (self.sum_losses(false_examples)
            + self.sum_losses(true_examples))
            self.display(3," split on",feat.__doc__,"has error",err,
            "splits
            into",len(true_examples),":",len(false_examples),"gamma="self.gamma)
            if err < best_error:
                best_feat = feat
                best_error=err
                best_partition = false_examples, true_examples
            self.display(2,"best split is on",best_feat.__doc__,
            "with err="best_error)
        return best_feat, best_partition

def sum_losses(self, data_subset):
    """returns sum of losses for dataset (with no more splits)
    There a single prediction for all leaves using leaf_prediction
    It is evaluated using split_to_optimize
    """
    prediction = self.leaf_value(data_subset, self.target.frange)
    error = sum(self.split_to_optimize(prediction, self.target(e))
        for e in data_subset)
    return error
def partition(data_subset, feature):
    """partitions the data_subset by the feature""
    true_examples = []
    false_examples = []
    for example in data_subset:
        if feature(example):
            true_examples.append(example)
        else:
            false_examples.append(example)
    return false_examples, true_examples

Test cases:

```python
from learnProblem import Data_set, Data_from_file

def testDT(data, print_tree=True, selections = None, **tree_args):
    """Prints errors and the trees for various evaluation criteria and ways to select leaves."
    if selections == None: # use selections suitable for target type
        selections = Predict.select[data.target.f_type]
    evaluation_criteria = Evaluate.all_criteria
    print("Split Choice","Leaf Choice\t","#leaves\t".join(ecrit.__doc__
        for ecrit in evaluation_criteria),sep="\t")
    for crit in evaluation_criteria:
        for leaf in selections:
            tree = DT_learner(data, split_to_optimize=crit,
                leaf_prediction=leaf,
                **tree_args).learn()
            print(crit.__doc__, leaf.__doc__, tree.num_leaves,
                "\t".join("{:7f}".format(data.evaluate_dataset(data.test,
                    tree, ecrit))
                    for ecrit in evaluation_criteria),sep="\t")
            if print_tree:
                print(tree.__doc__)

#DT_learner.max_display_level = 4
if __name__ == "__main__":
    # Choose one of the data files
    #data=Data_from_file('data/SPECT.csv', target_index=0);
    #print("SPECT.csv")
    #data=Data_from_file('data/iris.data', target_index=-1);
    #print("iris.data")
    data = Data_from_file('data/carbool.csv', target_index=-1)
    #data = Data_from_file('data/mail_reading.csv', target_index=-1);
    #print("mail_reading.csv")
    #data = Data_from_file('data/holiday.csv', num_train=19,
        target_index=-1); print("holiday.csv")
```
7.5 Cross Validation and Parameter Tuning

```python
testDT(data, print_tree=False)
```

Note that different runs may provide different values as they split the training and test sets differently. So if you have a hypothesis about what works better, make sure it is true for different runs.

Exercise 7.6 The current algorithm does not have a very sophisticated stopping criterion. What is the current stopping criterion? (Hint: you need to look at both `learn_tree` and `select_split`.)

Exercise 7.7 Extend the current algorithm to include in the stopping criterion

(a) A minimum child size; don’t use a split if one of the children has fewer elements that this.

(b) A depth-bound on the depth of the tree.

(c) An improvement bound such that a split is only carried out if error with the split is better than the error without the split by at least the improvement bound.

Which values for these parameters make the prediction errors on the test set the smallest? Try it on more than one dataset.

Exercise 7.8 Without any input features, it is often better to include a pseudo-count that is added to the counts from the training data. Modify the code so that it includes a pseudo-count for the predictions. When evaluating a split, including pseudo counts can make the split worse than no split. Does pruning with an improvement bound and pseudo-counts make the algorithm work better than with an improvement bound by itself?

Exercise 7.9 Some people have suggested using information gain (which is equivalent to greedy optimization of log loss) as the measure of improvement when building the tree, even in they want to have non-probabilistic predictions in the final tree. Does this work better than myopically choosing the split that is best for the evaluation criteria we will use to judge the final prediction?

7.5 Cross Validation and Parameter Tuning

To run the cross validation demo, in folder "aipython", load "learnCrossValidation.py", using e.g., `ipython -i learnCrossValidation.py`. Run the examples at the end to produce a graph like Figure 7.15. Note that different runs will produce different graphs, so your graph will not look like the one in the textbook. To try more examples, copy and paste the commented-out commands at the bottom of that file. This requires Python 3 with matplotlib.

The above decision tree overfits the data. One way to determine whether the prediction is overfitting is by cross validation. The code below implements k-fold cross validation, which can be used to choose the value of parameters
to best fit the training data. If we want to use parameter tuning to improve predictions on a particular data set, we can only use the training data (and not the test data) to tune the parameter.

In k-fold cross validation, we partition the training set into k approximately equal-sized folds (each fold is an enumeration of examples). For each fold, we train on the other examples, and determine the error of the prediction on that fold. For example, if there are 10 folds, we train on 90% of the data, and then test on remaining 10% of the data. We do this 10 times, so that each example gets used as a test set once, and in the training set 9 times.

The code below creates one copy of the data, and multiple views of the data. For each fold, fold enumerates the examples in the fold, and fold_complement enumerates the examples not in the fold.

```python
from learnProblem import Data_set, Data_from_file, Evaluate
from learnNoInputs import Predict
from learnDT import DT_learner
import matplotlib.pyplot as plt
import random

class K_fold_dataset(object):
    def __init__(self, training_set, num_folds):
        self.data = training_set.train.copy()
        self.target = training_set.target
        self.input_features = training_set.input_features
        self.num_folds = num_folds
        self.conditions = training_set.conditions

        random.shuffle(self.data)
        self.fold_boundaries = [(len(self.data)*i)//num_folds
                                for i in range(0,num_folds+1)]

    def fold(self, fold_num):
        for i in range(self.fold_boundaries[fold_num],
                        self.fold_boundaries[fold_num+1]):
            yield self.data[i]

    def fold_complement(self, fold_num):
        for i in range(0,self.fold_boundaries[fold_num]):
            yield self.data[i]
        for i in range(self.fold_boundaries[fold_num+1],len(self.data)):
            yield self.data[i]

The validation error is the average error for each example, where we test on each fold, and learn on the other folds.

```
for i in range(self.num_folds):
    predictor = learner(self,
                        train=list(self.fold_complement(i)),
                        **other_params).learn()
    error += sum(error_measure(predictor(e), self.target(e))
         for e in self.fold(i))
except ValueError:
    return float("inf")  # infinity
return error/len(self.data)

The plot_error method plots the average error as a function of the minimum number of examples in decision-tree search, both for the validation set and for the test set. The error on the validation set can be used to tune the parameter — choose the value of the parameter that minimizes the error. The error on the test set cannot be used to tune the parameters; if is were to be used this way then it cannot be used to test.

```python
52 def plot_error(data, criterion=Evaluate.squared_loss,
                leaf_prediction=Predict.empirical, num_folds=5, xscale='linear'):
    """Plots the error on the validation set and the test set
    with respect to settings of the minimum number of examples.
    xscale should be 'log' or 'linear'
    ""
    plt.ion()
    plt.xscale(xscale)  # change between log and linear scale
    plt.xlabel("min_child_weight")
    plt.ylabel("average " + criterion.__doc__")
    folded_data = K_fold_dataset(data, num_folds)
    verrors = []  # validation errors
    terrors = []  # test set errors
    for mcw in range(1, len(data.train)+2):
        verrors.append(folded_data.validation_error(DT_learner, criterion, leaf_prediction=leaf_prediction,
                                                  min_child_weight=mcw))
        tree = DT_learner(data, criterion, leaf_prediction=leaf_prediction,
                          min_child_weight=mcw).learn()
        terrors.append(data.evaluate_dataset(data.test, tree, criterion))
    plt.plot(range(1, len(data.train)+2), verrors, ls='-', color='k',
             label="validation for " + criterion.__doc__")
    plt.plot(range(1, len(data.train)+2), terrors, ls='--', color='r',
             label="test set for " + criterion.__doc__")
    plt.legend()
    plt.draw()  # The following produces Figure 7.15 of Poole and Mackworth [2017]
    # Different runs produce different plots, because folds change.
    # data = Data_from_file('data/SPECT.csv', target_index=0)
    # plot_error(data)  # warning, may take a long time depending on the
    # dataset
    # also try:
    http://aipython.org
```


7. Supervised Machine Learning

# data = Data_from_file('data/mail_reading.csv', target_index=-1)
# data = Data_from_file('data/carbool.csv', target_index=-1, seed=123)
# plot_error(data,
#     criterion=Evaluate.log_loss,leaf_prediction=Predict.laplace) # warning, may take a long time depending on the dataset

Note that different runs for the same data will have the same test error, but different validation error. If you rerun the Data_from_file, you will get the new test and training sets, and so the graph will change.

Exercise 7.10 Change the error plot so that it can evaluate the stopping criteria of the exercise of Section 7.6. Which criteria makes the most difference?

7.6 Linear Regression and Classification

Here we give a gradient descent searcher for linear regression and classification.

```python
from learnProblem import Learner
import random, math

class Linear_learner(Learner):
    def __init__(self, dataset, train=None,
                 learning_rate=0.1, max_init = 0.2,
                 squashed=True):
        """Creates a gradient descent searcher for a linear classifier.
        The main learning is carried out by learn()
        """
        self.dataset = dataset
        self.target = dataset.target
        if train==None:
            self.train = self.dataset.train
        else:
            self.train = train
        self.learning_rate = learning_rate
        self.squashed = squashed
        self.input_features = [one]+dataset.input_features # one is defined below
        self.weights = {feat:random.uniform(-max_init,max_init)
                        for feat in self.input_features}

    predictor predicts the value of an example from the current parameter settings.
    predictor.string gives a string representation of the predictor.
```

http://aipython.org  Version 0.9.4  March 2, 2022
7.6. Linear Regression and Classification

```python
def predictor(self, e):
    """returns the prediction of the learner on example e""
    linpred = sum(w*f(e) for f, w in self.weights.items())
    if self.squashed:
        return sigmoid(linpred)
    else:
        return linpred

def predictor_string(self, sig_dig=3):
    """returns the prediction function
    sig_dig is the number of significant digits in the numbers"
    doc = "".join(str(round(val, sig_dig))+"*"+feat.__doc__
    for feat, val in self.weights.items())
    if self.squashed:
        return "sigmoid(" + doc + ")"
    else:
        return doc

learn is the main algorithm of the learner. It does num_iter steps of stochastic
gradient descent with batch size = 1. The other parameters it gets from the
class.
```

```python
def learn(self, num_iter=100):
    for it in range(num_iter):
        self.display(2, "prediction=", self.predictor_string())
        for e in self.train:
            predicted = self.predictor(e)
            error = self.target(e) - predicted
            update = self.learning_rate*error
            for feat in self.weights:
                self.weights[feat] += update*feat(e)
        return self.predictor
```

`one` is a function that always returns 1. This is used for one of the input prop-
erties.

```python
def one(e):
    """1"
    return 1
```

`sigmoid(x)` is the function

\[
\frac{1}{1+e^{-x}}
\]

The inverse of `sigmoid` is the `logit` function

http://aipython.org

Version 0.9.4
March 2, 2022
def sigmoid(x):
    return 1/(1+math.exp(-x))

def logit(x):
    return -math.log(1/x-1)

sigmoid([x_0, v_2, ...]) returns [v_0, v_2, ...] where

\[ v_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \]

The inverse of sigmoid is the logit function

def softmax(xs, domain=None):
    """xs is a list of values, and
    domain is the domain (a list) or None if the list should be returned
    returns a distribution over the domain (a dict)
    """
    m = max(xs) # use of m prevents overflow (and all values underflowing)
    exps = [math.exp(x-m) for x in xs]
    s = sum(exps)
    if domain:
        return {d:v/s for (d,v) in zip(domain,exps)}
    else:
        return [v/s for v in exps]

def indicator(v, domain):
    return [1 if v==dv else 0 for dv in domain]

The following tests the learner on a data sets. Uncomment the other data
dsets for different examples.

from learnProblem import Data_set, Data_from_file, Evaluate
from learnProblem import Evaluate
import matplotlib.pyplot as plt

def test(**args):
    data = Data_from_file('data/SPECT.csv', target_index=0)
    # data = Data_from_file('data/mail_reading.csv', target_index=-1)
    # data = Data_from_file('data/carbool.csv', target_index=-1)
    learner = Linear_learner(data,**args)
    learner.learn()
    print("Function learned is", learner.predictor_string())
    for ecrit in Evaluate.all_criteria:
        test_error = data.evaluate_dataset(data.test, learner.predictor, ecrit)
        print(" Average", ecrit.__doc__, "is", test_error)

The following plots the errors on the training and test sets as a function of
the number of steps of gradient descent.
7.6. Linear Regression and Classification

```python
def plot_steps(learner=None, data = None, criterion=Evaluate.squared_loss, step=1, num_steps=1000, log_scale=True, legend_label=""):
    
    plots the training and test error for a learner. data is the learner_class is the class of the learning algorithm criterion gives the evaluation criterion plotted on the y-axis step specifies how many steps are run for each point on the plot num_steps is the number of points to plot

    if legend_label != ": legend_label=" "
    plt.ion()
    plt.xlabel("step")
    plt.ylabel("Average "+criterion.__doc__)
    if log_scale:
        plt.xscale('log') #plt.semilogx() #Makes a log scale
    else:
        plt.xscale('linear')
    if data is None:
        data = Data_from_file('data/holiday.csv', num_train=19, target_index=-1)
        #data = Data_from_file('data/SPECT.csv', target_index=0)
        #data = Data_from_file('data/mail_reading.csv', target_index=-1)
        #data = Data_from_file('data/carbool.csv', target_index=-1)
    #random.seed(None) # reset seed
    if learner is None:
        learner = Linear_learner(data)
    train_errors = []
    test_errors = []
    for i in range(1,num_steps+1,step):
        test_errors.append(data.evaluate_dataset(data.test, learner.predictor, criterion))
        train_errors.append(data.evaluate_dataset(data.train, learner.predictor, criterion))
        learner.display(2, "Train error:",train_errors[-1], "Test error:",test_errors[-1])
        learner.learn(num_iter=step)
    plt.plot(range(1,num_steps+1,step),train_errors,ls='-',label=legend_label+"training")
    plt.plot(range(1,num_steps+1,step),test_errors,ls='--',label=legend_label+"test")
    plt.legend()
    plt.draw()
    learner.display(1, "Train error:",train_errors[-1], "Test error:",test_errors[-1])
```

---

[http://aipython.org](http://aipython.org)  Version 0.9.4  March 2, 2022
if __name__ == "__main__":
    test()

# This generates the figure
# from learnProblem import Data_set_augmented, prod_feat
# data = Data_from_file('data/SPECT.csv', prob_test=0.5, target_index=0)
# dataplus = Data_set_augmented(data,[],[prod_feat])
# plot_steps(data=data,num_steps=1000)
# plot_steps(data=dataplus,num_steps=1000) # warning very slow

Exercise 7.11 The squashed learner only makes predictions in the range $(0, 1)$. If the output values are $\{1, 2, 3, 4\}$ there is no use prediction less than 1 or greater than 4. Change the squashed learner so that it can learn values in the range $(1, 4)$. Test it on the file 'data/car.csv'.

The following plots the prediction as a function of the number of steps of gradient descent. We first define a version of range that allows for real numbers (integers and floats).

```
def arange(start, stop, step):
    """returns enumeration of values in the range [start,stop) separated by step.
    like the built-in range(start,stop,step) but allows for integers and floats.
    Note that rounding errors are expected with real numbers. (or use numpy.arange)
    """
    while start<stop:
        yield start
        start += step

def plot_prediction(data,
    learner = None,
    minx = 0,
    maxx = 5,
    step_size = 0.01, # for plotting
    label = "function"):
    plt.ion()
    plt.xlabel("x")
    plt.ylabel("y")
    if learner is None:
        learner = Linear_learner(data, squashed=False)
    learner.learning_rate=0.001
    learner.learn(100)
    learner.learning_rate=0.0001
    learner.learn(1000)
    learner.learning_rate=0.00001
    learner.learn(10000)
    learner.display(1,"function learned is", learner.predictor_string()),
```
7.6. Linear Regression and Classification

from learnProblem import Data_set_augmented, power_feat

def plot_polynomials(data, learner_class = Linear_learner, max_degree = 5, minx = 0, maxx = 5, num_iter = 100000, learning_rate = 0.0001, step_size = 0.01, # for plotting):
    plt.ion()
    plt.xlabel("x")
    plt.ylabel("y")
    pl.plot([e[0] for e in data.train],[e[-1] for e in data.train],"ko",label="data")
    x_values = list(arange(minx,maxx,step_size))
    line_styles = [ '-', '--', '-.', ':', ' ']
    colors = [ '0.5', 'k', 'k', 'k', 'k']
    for degree in range(max_degree):
        data_aug = Data_set_augmented(data,[power_feat(n) for n in range(1,degree+1)], include_orig=False)
        learner = learner_class(data_aug,squashed=False)
        learner.learning_rate = learning_rate
        learner.learn(num_iter)
        learner.display(1,"For degree",degree,"function learned is", learner.predictor_string(), "error=",data.evaluate_dataset(data.train, learner.predictor, Evaluate.squared_loss))
        ls = line_styles[degree % len(line_styles)]
        col = colors[degree % len(colors)]
        pl.plot(x_values,[learner.predictor([x]) for x in x_values], linestyle=ls, color=col, label="degree"+str(degree))
        pl.legend(loc='upper left')
        pl.draw()
7. Supervised Machine Learning

```python
# data0 = Data_from_file('data/simp_regr.csv', prob_test=0,  
  boolean_features=False, target_index=-1)
# plot_prediction(data0)
# plot_polynomials(data0)
# datam = Data_from_file('data/mail_reading.csv', target_index=-1)
# plot_prediction(datam)

7.6.1 Batched Stochastic Gradient Descent

This implements batched stochastic gradient descent. If the batch size is 1, it
 can be simplified by not storing the differences in d, but applying them directly;
 this would be equivalent to the original code!

This overrides the learner LinearLearner. Note that the comparison with
regular gradient descent is unfair as the number of updates per step is not the
same. (How could it have made more fair?)

```

```
7.7. Boosting

The following code implements functional gradient boosting for regression.

A Boosted dataset is created from a base dataset by subtracting the prediction of the offset function from each example. This does not save the new dataset, but generates it as needed. The amount of space used is constant, independent on the size of the data set.

```python
from learnProblem import Data_set, Learner, Evaluate
from learnNoInputs import Predict
from learnLinear import sigmoid
import statistics
import random

class Boosted_dataset(Data_set):
    def __init__(self, base_dataset, offset_fun, subsample=1.0):
        
        self.base_dataset = base_dataset
        self.offset_fun = offset_fun
        self.train = random.sample(base_dataset.train, int(subsample*len(base_dataset.train)))
        self.test = base_dataset.test
        
        self.input_features = self.base_dataset.input_features
        def newout(e):
            return self.base_dataset.target(e) - self.offset_fun(e)
        newout.frange = self.base_dataset.target.frange
        newout.ftype = self.infer_type(newout.frange)
        self.target = newout

        def conditions(self, *args, colsample_bytree=0.5, **nargs):
            conds = self.base_dataset.conditions(*args, **nargs) +
            return random.sample(conds, int(colsample_bytree*len(conds)))

    def create_features(self):
        defines a new target
        self.input_features = self.base_dataset.input_features
        def newout(e):
            return self.base_dataset.target(e) - self.offset_fun(e)
        newout.frange = self.base_dataset.target.frange
        newout.ftype = self.infer_type(newout.frange)
        self.target = newout

        def conditions(self, *args, colsample_bytree=0.5, **nargs):
            conds = self.base_dataset.conditions(*args, **nargs)
            return random.sample(conds, int(colsample_bytree*len(conds)))
```

A boosting learner takes in a dataset and a base learner, and returns a new predictor. The base learner, takes a dataset, and returns a Learner object.

http://aipython.org

Version 0.9.4
March 2, 2022
class Boosting_learner(Learner):
    def __init__(self, dataset, base_learner_class, subsample=0.8):
        self.dataset = dataset
        self.base_learner_class = base_learner_class
        self.subsample = subsample
        mean = sum(self.dataset.target(e) for e in self.dataset.train) / len(self.dataset.train)
        self.predictor = lambda e: mean  # function that returns mean for each example
        self.predictor.__doc__ = "lambda e:" + str(mean)
        self.offsets = [self.predictor]  # list of base learners
        self.predictors = [self.predictor]  # list of predictors
        self.errors = [data.evaluate_dataset(data.test, self.predictor, Evaluate.squared_loss)]
        self.display(1, "Predict mean test set mean squared loss=",
        self.errors[0])

    def learn(self, num_ensembles=10):
        """
        adds num_ensemble learners to the ensemble.
        returns a new predictor.
        """
        for i in range(num_ensembles):
            train_subset = Boosted_dataset(self.dataset, self.predictor, subsample=self.subsample)
            learner = self.base_learner_class(train_subset)
            new_offset = learner.learn()
            self.offsets.append(new_offset)
            def new_pred(e, old_pred=self.predictor, off=new_offset):
                return old_pred(e) + off(e)
            self.predictor = new_pred
            self.predictors.append(new_pred)
            self.errors.append(data.evaluate_dataset(data.test, self.predictor, Evaluate.squared_loss))
            self.display(1, f"Iteration {len(self.offsets)-1},treesize = {new_offset.num_leaves}. mean squared loss={self.errors[-1]}")
        return self.predictor

For testing, sp_DT_learner returns a learner that predicts the mean at the leaves and is evaluated using squared loss. It can also take arguments to change the default arguments for the trees.
7.7. Boosting

```python
leaf_prediction=Predict.mean,**nargs):
    """Creates a learner with different default arguments replaced by
    **nargs
    """

def new_learner(dataset):
    return DT_learner(dataset, split_to_optimize=split_to_optimize,
    leaf_prediction=leaf_prediction,**nargs)

return new_learner

#data = Data_from_file('data/car.csv', target_index=-1) regression
#data = Data_from_file('data/student/student-mat-nq.csv', separator=';', has_header=True, target_index=-1, include_only=list(range(30)) + [32]) #2.0537973790924946
#data = Data_from_file('data/SPECT.csv', target_index=0, seed=62) #123)
#data = Data_from_file('data/mail_reading.csv', target_index=-1)
#data = Data_from_file('data/holiday.csv', num_train=19, target_index=-1)
#learner10 = Boosting_learner(data,
#    sp_DT_learner(split_to_optimize=Evaluate.squared_loss,
#    leaf_prediction=Predict.mean, min_child_weight=10))
#learner7 = Boosting_learner(data, sp_DT_learner(0.7))
#learner5 = Boosting_learner(data, sp_DT_learner(0.5))
#predictor9 = learner9.learn(10)
#for i in learner9.offsets: print(i.__doc__)

import matplotlib.pyplot as plt

def plot_boosting_trees(data, steps=10, mcws=[30,20,20,10], gammas=[100,200,300,500]):
    # to reduce clutter uncomment one of following two lines
    #mcws=[10]
    #gammas=[200]
    learners = [(mcw, gamma, Boosting_learner(data,
        sp_DT_learner(min_child_weight=mcw, gamma=gamma)))
        for gamma in gammas for mcw in mcws]

plt.ion()
plt.xscale('linear') # change between log and linear scale
plt.xlabel("number of trees")
plt.ylabel("mean squared loss")
markers = (m+c for c in ['k','g','r','b','m','c','y'] for m in ['-','--','-',':'])
for (mcw,gamma,learner) in learners:
    data.display(1,"min_child_weight={mcw}, gamma={gamma}"")
    learner.learn(steps)
    plt.plot(range(steps+1), learner.errors, next(markers), label=f"min_child_weight={mcw}, gamma={gamma}"")
plt.legend()
plt.draw()

# plot_boosting_trees(data)
```

http://aipython.org
160

7. Supervised Machine Learning

7.7.1 Gradient Tree Boosting
The following implements gradient Boosted trees for classification. If you want
to use this gradient tree boosting for a real problem, we recommend using
XGBoost [Chen and Guestrin, 2016].
GTB_learner subclasses DT-learner. The method learn_tree is used unchanged. DT-learner assumes that the value at the leaf is the prediction of the
leaf, thus leaf_value needs to be overridden. It also assumes that all nodes
at a leaf have the same prediction, but in GBT the elements of a leaf can have
different values, depending on the previous trees. Thus sum_losses also needs
to be overridden.
learnBoosting.py — (continued)
124
125
126
127
128
129
130

class GTB_learner(DT_learner):
def __init__(self, dataset, number_trees, lambda_reg=1, gamma=0,
**dtargs):
DT_learner.__init__(self, dataset,
split_to_optimize=Evaluate.log_loss, **dtargs)
self.number_trees = number_trees
self.lambda_reg = lambda_reg
self.gamma = gamma
self.trees = []

131
132
133
134

135
136
137
138
139
140

def learn(self):
for i in range(self.number_trees):
tree =
self.learn_tree(self.dataset.conditions(self.max_num_cuts),
self.train)
self.trees.append(tree)
self.display(1,f"""Iteration {i} treesize = {tree.num_leaves}
train logloss={
self.dataset.evaluate_dataset(self.dataset.train,
self.gtb_predictor, Evaluate.log_loss)
} test logloss={
self.dataset.evaluate_dataset(self.dataset.test,
self.gtb_predictor, Evaluate.log_loss)}""")
return self.gtb_predictor

141
142
143
144
145
146

def gtb_predictor(self, example, extra=0):
"""prediction for example,
extras is an extra contribution for this example being considered
"""
return sigmoid(sum(t(example) for t in self.trees)+extra)

147
148
149
150
151
152

def leaf_value(self, egs, domain=[0,1]):
"""value at the leaves for examples egs
domain argument is ignored"""
pred_acts = [(self.gtb_predictor(e),self.target(e)) for e in egs]
return sum(a-p for (p,a) in pred_acts) /(sum(p*(1-p) for (p,a) in
pred_acts)+self.lambda_reg)

http://aipython.org

Version 0.9.4

March 2, 2022


```python
def sum_losses(self, data_subset):
    """returns sum of losses for dataset (assuming a leaf is formed
    with no more splits)
    """
    leaf_val = self.leaf_value(data_subset)
    error = sum(Evaluate.log_loss(self.gtb_predictor(e, leaf_val),
                                self.target(e))
                for e in data_subset) + self.gamma
    return error
```

Testing

```python
# data = Data_from_file('data/carbool.csv', target_index=-1, seed=123)
# gtb_learner = GTB_learner(data, 10)
# gtb_learner.learn()
```
Neural Networks and Deep Learning

Warning: this is not meant to be an efficient implementation of deep learning. If you want to do serious machine learning on medium-sized or large data, we would recommend Keras (https://keras.io) [Chollet, 2021] or PyTorch (https://pytorch.org), which are very efficient, particularly on GPUs. They are, however, black boxes. The AIPython neural network code should be seen like a car engine made of glass; you can see exactly how it works, even if it is not fast.

We have made parameters that are the same as in Keras have the same names.

8.1 Layers

A neural network is built from layers. This provides a modular implementation of layers. Layers can easily be stacked in many configurations. A layer needs to implement a function to compute the output values from the inputs, a way to back-propagate the error, and perhaps update its parameters.

```python
from learnProblem import Learner, Data_set, Data_from_file, Data_from_files, Evaluate
from learnLinear import sigmoid, one, softmax, indicator
import random, math, time

class Layer(object):
    def __init__(self, nn, num_outputs=None):
```
"""Given a list of inputs, outputs will produce a list of length num_outputs.
n is the neural network this layer is part of
num outputs is the number of outputs for this layer.
"""
self.nn = nn
self.num_inputs = nn.num_outputs # output of nn is the input to
this layer
if num_outputs:
    self.num_outputs = num_outputs
else:
    self.num_outputs = nn.num_outputs # same as the inputs
def output_values(self,input_values):
    """Return the outputs for this layer for the given input values.
in is a list of the inputs to this layer (of length num_inputs)
returns a list of length self.num_outputs
"""
raise NotImplemented("output_values") # abstract method
def backprop(self,errors):
    """Backpropagate the errors on the outputs
errors is a list of errors for the outputs (of length self.num_outputs).
Returns the errors for the inputs to this layer (of length self.num_inputs).
You can assume that this is only called after corresponding
output_values,
which can remember information information required for the
back-propagation.
"""
raise NotImplemented("backprop") # abstract method
def update(self):
    """Updates parameters after a batch.
overridden by layers that have parameters
""
pass
A linear layer maintains an array of weights. self.weights[o][i] is the weight
between input i and output o. A 1 is added to the end of the inputs. The default
initialization is the Glorot uniform initializer [Glorot and Bengio, 2010], which
is the default in Keras. An alternative is to provide a limit, in which case the
values are selected uniformly in the range [−limit,limit]. Keras treats the bias
separately, and defaults to zero.
def __init__(self, nn, num_outputs, limit=None):
    """A completely connected linear layer.
    nn is a neural network that the inputs come from
    num_outputs is the number of outputs
    the random initialization of parameters is in range [-limit,limit]
    """
    Layer.__init__(self, nn, num_outputs)
    if limit is None:
        limit =math.sqrt(6/(self.num_inputs+self.num_outputs))
    # self.weights[o][i] is the weight between input i and output o
    self.weights = [[random.uniform(-limit, limit) if inf <
        self.num_inputs else 0
        for inf in range(self.num_inputs+1)]
        for outf in range(self.num_outputs)]
    self.delta = [[0 for inf in range(self.num_inputs+1)]
        for outf in range(self.num_outputs)]

def output_values(self,input_values):
    """Returns the outputs for the input values.
    It remembers the values for the backprop.
    """
    Note in self.weights there is a weight list for every output,
    so wts in self.weights loops over the outputs.
    The bias is the *last* value of each list in self.weights.
    """
    self.inputs = input_values + [1]
    return [sum(w*val for (w,val) in zip(wts,self.inputs))
        for wts in self.weights]

def backprop(self,errors):
    """Backpropagate the errors, updating the weights and returning the
    error in its inputs.
    """
    input_errors = [0]*(self.num_inputs+1)
    for out in range(self.num_outputs):
        for inp in range(self.num_inputs+1):
            input_errors[inp] += self.weights[out][inp] * errors[out]
            self.delta[out][inp] += self.inputs[inp] * errors[out]
    return input_errors[:-1] # remove the error for the "1"

def update(self):
    """Updates parameters after a batch"
    batch_step_size = self.nn.learning_rate / self.nn.batch_size
    for out in range(self.num_outputs):
        for inp in range(self.num_inputs+1):
            self.weights[out][inp] -= batch_step_size *
                self.delta[out][inp]
            self.delta[out][inp] = 0

The standard activation function for hidden nodes is the **ReLU**.
8. Neural Networks and Deep Learning

8.2 Feedforward Networks

class ReLU_layer(Layer):
    """Rectified linear unit (ReLU) \( f(z) = \max(0, z) \).
    The number of outputs is equal to the number of inputs.
    """
    def __init__(self, nn):
        Layer.__init__(self, nn)

    def output_values(self,input_values):
        """Returns the outputs for the input values.
        It remembers the input values for the backprop.
        """
        self.input_values = input_values
        self.outputs= [\max(0,inp) for inp in input_values]
        return self.outputs

    def backprop(self,errors):
        """Returns the derivative of the errors"""
        return [e if inp>0 else 0 for e,inp in zip(errors, self.input_values)]

One of the old standards for the activation function for hidden layers is the sigmoid. It is included here to experiment with.

class Sigmoid_layer(Layer):
    """Sigmoids of the inputs.
    The number of outputs is equal to the number of inputs.
    Each output is the sigmoid of its corresponding input.
    """
    def __init__(self, nn):
        Layer.__init__(self, nn)

    def output_values(self,input_values):
        """Returns the outputs for the input values.
        It remembers the output values for the backprop.
        """
        self.outputs= [sigmoid(inp) for inp in input_values]
        return self.outputs

    def backprop(self,errors):
        """Returns the derivative of the errors"""
        return [e*out*(1-out) for e,out in zip(errors, self.outputs)]
8.2. Feedforward Networks

```python
def __init__(self, dataset, validation_proportion = 0.1,
             learning_rate=0.001):
    """Creates a neural network for a dataset,
layers is the list of layers
    """
    self.dataset = dataset
    self.output_type = dataset.target.ftype
    self.learning_rate = learning_rate
    self.input_features = dataset.input_features
    self.num_outputs = len(self.input_features)
    validation_num = int(len(self.dataset.train)*validation_proportion)
    if validation_num > 0:
        random.shuffle(self.dataset.train)
        self.validation_set = self.dataset.train[-validation_num:]
        self.training_set = self.dataset.train[:-validation_num]
    else:
        self.validation_set = []
        self.training_set = self.dataset.train
    self.layers = []
    self.bn = 0 # number of batches run

def add_layer(self,layer):
    """add a layer to the network.
    Each layer gets number of inputs from the previous layers outputs.
    """
    self.layers.append(layer)
    self.num_outputs = layer.num_outputs

def predictor(self,ex):
    """Predicts the value of the first output for example ex.
    """
    values = [f(ex) for f in self.input_features]
    for layer in self.layers:
        values = layer.output_values(values)
    return sigmoid(values[0]) if self.output_type == "boolean"
    else softmax(values, self.dataset.target.frange) if
    self.output_type == "categorical"
    else values[0]

def predictor_string(self):
    return "not implemented"

The learn method learns a network.

```
overrides epochs if provided (allows for fractions of epochs)

self.batch_size = min(batch_size, len(self.training_set))  # don't have batches bigger than training size

if num_iter is None:
    num_iter = (epochs * len(self.training_set)) // self.batch_size
#self.display(0,"Batch","\t".join(criterion.__doc__ for criterion in Evaluate.all_criteria))

for i in range(num_iter):
    batch = random.sample(self.training_set, self.batch_size)
    for e in batch:
        # compute all outputs
        values = [f(e) for f in self.input_features]
        for layer in self.layers:
            values = layer.output_values(values)
            # backpropagate
            predicted = [sigmoid(v) for v in values] if self.output_type == "boolean"
            else softmax(values) if self.output_type == "categorical"
            else values
            actuals = indicator(self.dataset.target(e), self.dataset.target.frange) \
            if self.output_type == "categorical"
            else [self.dataset.target(e)]
            errors = [pred-obsd for (obsd,pred) in zip(actuals,predicted)]
            for layer in reversed(self.layers):
                errors = layer.backprop(errors)
        # Update all parameters in batch
        for layer in self.layers:
            layer.update()
        self.bn+=1
    self.display(0,self.bn,"	",
        "\tt\t".join("{:.4f}".format(
            self.dataset.evaluate_dataset(self.validation_set, self.predictor, criterion))
            for criterion in Evaluate.all_criteria), sep="")

8.3 Improved Optimization

8.3.1 Momentum

... (continued)...

class Linear_complete_layer_momentum(Linear_complete_layer):
    """A completely connected layer""
    def __init__(self, nn, num_outputs, limit=None, alpha=0.9, epsilon = 1e-07, vel0=0):
        """A completely connected linear layer.

http://aipython.org
8.3. Improved Optimization

nn is a neural network that the inputs come from
num_outputs is the number of outputs
max_init is the maximum value for random initialization of
parameters
vel0 is the initial velocity for each parameter

```
Linear_complete_layer.__init__(self, nn, num_outputs, limit=limit)
# self.weights[o][i] is the weight between input i and output o
self.velocity = [[vel0 for inf in range(self.num_inputs+1)]
    for outf in range(self.num_outputs)]
self.alpha = alpha
self.epsilon = epsilon
```

def update(self):
    """updates parameters after a batch""
    batch_step_size = self.nn.learning_rate / self.nn.batch_size
    for out in range(self.num_outputs):
        for inp in range(self.num_inputs+1):
            self.velocity[out][inp] = self.alpha*self.velocity[out][inp]
            self.weights[out][inp] += self.velocity[out][inp]
            self.delta[out][inp] = 0

8.3.2 RMS-Prop

```
class Linear_complete_layer_RMS_Prop(Linear_complete_layer):
    """a completely connected layer""
    def __init__(self, nn, num_outputs, limit=None, rho=0.9, epsilon = 1e-07):
        """A completely connected linear layer.
        nn is a neural network that the inputs come from
        num_outputs is the number of outputs
        max_init is the maximum value for random initialization of
        parameters"
        Linear_complete_layer.__init__(self, nn, num_outputs, limit=limit)
        # self.weights[o][i] is the weight between input i and output o
        self.ms = [[0 for inf in range(self.num_inputs+1)]
                    for outf in range(self.num_outputs)]
        self.rho = rho
        self.epsilon = epsilon

    def update(self):
        """updates parameters after a batch""
        for out in range(self.num_outputs):
            for inp in range(self.num_inputs+1):
                gradient = self.delta[out][inp] / self.nn.batch_size
                self.ms[out][inp] = self.rho*self.ms[out][inp] + (1-self.rho) *
                    gradient**2
```

http://aipython.org  Version 0.9.4  March 2, 2022
self.weights[out][inp] -=
    self.nn.learning_rate/(self.ms[out][inp]+self.epsilon)**0.5
    * gradient
self.delta[out][inp] = 0

8.3.3 Examples

The following constructs a neural network with one hidden layer. The hidden
layer has width 2 with a ReLU activation function. The output layer used a
sigmoid

```python
#data = Data_from_file('data/mail_reading.csv', target_index=-1)
data = Data_from_file('data/mail_reading_consis.csv', target_index=-1)
data = Data_from_file('data/SPECT.csv', prob_test=0.5, target_index=0)
data = Data_from_file('data/iris.data', prob_test=0.2, target_index=-1) #
    150 examples approx 120 test + 30 test
#data = Data_from_file('data/if_x_then_y_else_z.csv', num_train=8,
    target_index=-1) # not linearly sep
#data = Data_from_file('data/holiday.csv', target_index=-1) #,
    num_train=19)
#data = Data_from_file('data/processed.cleveland.data', target_index=-1)
random.seed(None)

nn1 = NN(data)
nn1.add_layer(Linear_complete_layer(nn1,3))
#nn1.add_layer(Sigmoid_layer(nn1)) # comment this or the next
nn1.add_layer(ReLU_layer(nn1))
#nn1.add_layer(Linear_complete_layer(nn1,1)) # when using
    output_type="boolean"
nn1.add_layer(Linear_complete_layer(nn1,1)) # when using
    output_type="categorical"
#nn1.learn(epochs = 100)

nn_r1 = NN(data)
nn_r1.add_layer(Linear_complete_layer_RMS_Prop(nn_r1,3))
#nn_r1.add_layer(Sigmoid_layer(nn_r1)) # comment this or the next
nn_r1.add_layer(ReLU_layer(nn_r1))
#nn_r1.add_layer(Linear_complete_layer(nn_r1,1)) # when using
    output_type="boolean"
nn_r1.add_layer(Linear_complete_layer_RMS_Prop(nn_r1,1)) # when using
    output_type="categorical"
#nn_r1.learn(epochs = 100)

nnm1 = NN(data)
nnm1.add_layer(Linear_complete_layer_momentum(nnm1,3))
#nnm1.add_layer(Sigmoid_layer(nnm1)) # comment this or the next
nnm1.add_layer(ReLU_layer(nnm1))
#nnm1.add_layer(Linear_complete_layer(nnm1,1)) # when using
    output_type="boolean"

http://aipython.org  Version 0.9.4  March 2, 2022
8.3. Improved Optimization

nnm1.add_layer(Linear_complete_layer_momentum(nnm1,1)) # when using output_type="categorical"
#nnm1.learn(epochs = 100)

nn2 = NN(data) #"boolean") #
nn2.add_layer(Linear_complete_layer_RMS_Prop(nn2,2))
nn2.add_layer(ReLU_layer(nn2))
nn2.add_layer(Linear_complete_layer_RMS_Prop(nn2,1)) # when using output_type="categorical"

nn3 = NN(data) #"boolean") #
nn3.add_layer(Linear_complete_layer_RMS_Prop(nn3,5))
nn3.add_layer(ReLU_layer(nn3))
nn3.add_layer(Linear_complete_layer_RMS_Prop(nn3,1)) # when using output_type="categorical"

nn0 = NN(data,learning_rate=0.05)
nn0.add_layer(Linear_complete_layer(nn0,1)) # categorical linear regression
#nn0.add_layer(Linear_complete_layer_RMS_Prop(nn0,1)) # categorical linear regression

Plotting.

```python
from learnLinear import plot_steps
from learnProblem import Evaluate

# To show plots:
# plot_steps(learner = nn1, data = data, criterion=Evaluate.log_loss, num_steps=10000, log_scale=False, legend_label="nn1")
# plot_steps(learner = nn2, data = data, criterion=Evaluate.log_loss, num_steps=10000, log_scale=False, legend_label="nn2")
# plot_steps(learner = nn3, data = data, criterion=Evaluate.log_loss, num_steps=100000, log_scale=False, legend_label="nn3")
# plot_steps(learner = nn0, data = data, criterion=Evaluate.log_loss, num_steps=10000, log_scale=False, legend_label="nn0")
# plot_steps(learner = nn0, data = data, criterion=Evaluate.accuracy, num_steps=10000, log_scale=False, legend_label="nn0")
# plot_steps(learner = nn1, data = data, criterion=Evaluate.accuracy, num_steps=10000, log_scale=False, legend_label="nn1")
# plot_steps(learner = nn2, data = data, criterion=Evaluate.accuracy, num_steps=10000, log_scale=False, legend_label="nn2")
# plot_steps(learner = nn3, data = data, criterion=Evaluate.accuracy, num_steps=10000, log_scale=False, legend_label="nn3")

# Print some training examples
#for eg in random.sample(data.train,10): print(eg,nn1.predictor(eg))
```

http://aipython.org
# Print some test examples
#for eg in random.sample(data.test,10): print(eg,nn1.predictor(eg))

# To see the weights learned in linear layers
# nn1.layers[0].weights
# nn1.layers[2].weights

# Print test:
# for e in data.train: print(e,nn0.predictor(e))

def test(data, hidden_widths = [5], epochs=100,
optimizers = [Linear_complete_layer,
              Linear_complete_layer_momentum,
              Linear_complete_layer_RMS_Prop]):
    data.display(0,"Batch\t","\t".join(criterion.__doc__
    Evaluate.all_criteria))
    for optimizer in optimizers:
        nn = NN(data)
        for width in hidden_widths:
            nn.add_layer(optimizer(nn,width))
            nn.add_layer(ReLU_layer(nn))
        if data.target.ftype == "boolean":
            nn.add_layer(optimizer(nn,1))
        else:
            error(f"Not implemented: {data.output_type}"
        nn.learn(epochs)

The following tests on MNIST. The original files are from http://yann.lecun.
com/exdb/mnist/. This code assumes you use the csv files from https://pjreddie.
com/projects/mnist-in-csv/, and put them in the directory ../MNIST/. Note
that this is very inefficient; you would be better to use Keras or Pytorch. There
are 28 * 28 = input units and 512 hidden units, which makes 401,408 parame-
ters for the lowest linear layer. So don’t be surprised when it takes many hours.

# Simplified version: (6000 training instances)
# data_mnist = Data_from_file('..../MNIST/mnist_train.csv', prob_test=0.9,
#     target_index=0, boolean_features=False, target_type="categorical")

# Full version:
# data_mnist = Data_from_files('..../MNIST/mnist_train.csv',
#     '..../MNIST/mnist_test.csv', target_index=0, boolean_features=False,
#     target_type="categorical")

# nn_mnist = NN(data_mnist, validation_proportion = 0.02,
#     learning_rate=0.001) #validation set = 1200
# nn_mnist.add_layer(Linear_complete_layer_RMS_Prop(nn_mnist,512));
# nn_mnist.add_layer(ReLU_layer(nn_mnist));
# nn_mnist.add_layer(Linear_complete_layer_RMS_Prop(nn_mnist,10))
# start_time = time.perf_counter();nn_mnist.learn(epochs=1,
#     batch_size=128);end_time = time.perf_counter();print("Time:", end_time

http://aipython.org Version 0.9.4 March 2, 2022
8.3. Improved Optimization

```python
- start_time,"seconds") #1 epoch
# determine test error:
data_mnist.evaluate_dataset(data_mnist.test, nn_mnist.predictor,
    Evaluate.accuracy)
# Print some random predictions:
for eg in random.sample(data_mnist.test,10):
    print(data_mnist.target(eg),nn_mnist.predictor(eg),nn_mnist.predictor(eg)[data_mnist.target(eg)])
```

**Exercise 8.1** In the definition of nn1 above, for each of the following, first hypothesize what will happen, then test your hypothesis, then explain whether you testing confirms your hypothesis or not. Test it for more than one data set, and use more than one run for each data set.

(a) Which fits the data better, having a sigmoid layer or a ReLU layer after the first linear layer?

(b) Which is faster, having a sigmoid layer or a ReLU layer after the first linear layer?

(c) What happens if you have both the sigmoid layer and then a ReLU layer after the first linear layer and before the second linear layer?

(d) What happens if you have neither the sigmoid layer nor a ReLU layer after the first linear layer?

(e) What happens if you have a ReLU layer then a sigmoid layer after the first linear layer and before the second linear layer?

**Exercise 8.2** Do some
Chapter 9

Reasoning Under Uncertainty

9.1 Representing Probabilistic Models

A variable consists of a name, a domain and an optional (x,y) position (for displaying). The domain of a variable is a list or a tuple, as the ordering will matter in the representation of factors.

```python
import random

class Variable(object):
    """A random variable.
    name (string) - name of the variable
domain (list) - a list of the values for the variable.
Variables are ordered according to their name.
"""

    def __init__(self, name, domain, position=None):
        """Variable
domain a list of printable values
position of form (x,y)
"""
        self.name = name # string
        self.domain = domain # list of values
        self.position = position if position else (random.random(),
        random.random())
        self.size = len(domain)

    def __str__(self):
        return self.name
```

175
9. Reasoning Under Uncertainty

9.2 Representing Factors

A factor is, mathematically, a function from variables into a number; that is given a value for each of its variable, it gives a number. Factors are used for conditional probabilities, utilities in the next chapter, and are explicitly constructed by some algorithms (in particular variable elimination).

A variable assignment, or just assignment, is represented as a variable : value dictionary. A factor can be evaluated when all of its variables are assigned. The method get_value evaluates the factor for an assignment. The assignment can include extra variables not in the factor. This method needs to be defined for every subclass.

```python
def __str__(self):
    return self.name # f"Variable({self.name})"
```

The method __str__ returns a brief definition (like “f7(X,Y,Z)”). The method to_table returns string representations of a table showing all of the assignments of values to variables, and the corresponding value.

```python
def __str__(self):
    return
```
9.3 Conditional Probability Distributions

A conditional probability distribution (CPD) is a type of factor that represents a conditional probability. A CPD representing $P(X \mid Y_1 \ldots Y_k)$ is a type of factor, where given values for $X$ and each $Y_i$ returns a number.

```python
class CPD(Factor):
    __init__(self, child, parents):
        """represents P(variable | parents)"""
        self.parents = parents
        self.child = child
        Factor.__init__(self, parents+[child])

    __str__(self):
        """A brief description of a factor using in tracing"""
        if self.parents:
            return f"P({self.child}|{', '.join(str(p) for p in self.parents))}"
```

http://aipython.org

Version 0.9.4 March 2, 2022
else:
    return f"P({self.child})"

__repr__ = __str__

The simplest CPD is the constant that has probability 1 when the child has the value specified.

class ConstantCPD(CPD):
    def __init__(self, variable, value):
        CPD.__init__(self, variable, [])
        self.value = value
    def get_value(self, assignment):
        return 1 if self.value==assignment[self.child] else 0

9.3.1 Logistic Regression

A logistic regression CPD, for Boolean variable \(X\) represents \(P(X=\text{True} \mid Y_1 \ldots Y_k)\), using \(k+1\) real-values weights so

\[
P(X=\text{True} \mid Y_1 \ldots Y_k) = \text{sigmoid}(w_0 + \sum_i w_i Y_i)
\]

where for Boolean \(Y_i\), True is represented as 1 and False as 0.

class LogisticRegression(CPD):
    def __init__(self, child, parents, weights):
        """A logistic regression representation of a conditional probability. child is the Boolean (or 0/1) variable whose CPD is being defined parents is the list of parents weights is list of parameters, such that weights[i+1] is the weight for parents[i] """
        assert len(weights) == 1+len(parents)
        CPD.__init__(self, child, parents)
        self.weights = weights
    def get_value(self, assignment):
        assert self.can_evaluate(assignment)
        prob = sigmoid(self.weights[0])
        for i in range(len(self.parents))
            prob = prob + sum(self.weights[i+1]*assignment[self.parents[i]])
        if assignment[self.child]: #child is true
            return prob
        return (1-prob)
9.3.2 Noisy-or

A noisy-or, for Boolean variable $X$ with Boolean parents $Y_1 \ldots Y_k$ is parametrized by $k + 1$ parameters $p_0, p_1, \ldots, p_k$, where each $0 \leq p_i \leq 1$. The semantics is defined as though there are $k + 1$ hidden variables $Z_0, Z_1 \ldots Z_k$, where $P(Z_0) = p_0$ and $P(Z_i \mid Y_i) = p_i$ for $i \geq 1$, and where $X$ is true if and only if $Z_0 \lor Z_1 \lor \cdots \lor Z_k$ (where $\lor$ is “or”). Thus $X$ is false if all of the $Z_i$ are false. Intuitively, $Z_0$ is the probability of $X$ when all $Y_i$ are false and each $Z_i$ is a noisy (probabilistic) measure that $Y_i$ makes $X$ true, and $X$ only needs one to make it true.

```python
class NoisyOR(CPD):
    def __init__(self, child, parents, weights):
        """A noisy representation of a conditional probability.
        variable is the Boolean (or 0/1) child variable whose CPD is being defined
        parents is the list of Boolean (or 0/1) parents
        weights is list of parameters, such that weights[i+1] is the weight for parents[i]
        ""
        assert len(weights) == 1+len(parents)
        CPD.__init__(self, child, parents)
        self.weights = weights

    def get_value(self, assignment):
        assert self.can_evaluate(assignment)
        probfalse = (1-self.weights[0])*math.prod(1-self.weights[i+1]
                    for i in range(len(self.parents))
                    if assignment[self.parents[i]])

        if assignment[self.child]:
            return 1-probfalse
        else:
            return probfalse
```

9.3.3 Tabular Factors

A tabular factor is a factor that represents each assignment of values to variables separately. It is represented by a Python array (or python dict). If the variables are $V_1, V_2, \ldots, V_k$, the value of $f(V_1 = v_1, V_2 = v_1, \ldots, V_k = v_k)$ is stored in $f[v_1][v_2][\ldots][v_k]$.

If the domain of $V_i$ is $[0, \ldots, n_i - 1]$ this can be represented as an array. Otherwise we can use a dictionary. Python is nice in that it doesn’t care, whether an array or dict is used except when enumerating the values; enumerating a dict gives the keys (the variables) but enumerating an array gives the values. So we have to be careful not to do this.
from functools import reduce

class TabFactor(Factor):
    def __init__(self, variables, values):
        Factor.__init__(self, variables)
        self.values = values

    def get_value(self, assignment):
        return self.get_val_rec(self.values, self.variables, assignment)

    def get_val_rec(self, value, variables, assignment):
        if variables == []:
            return value
        else:
            return self.get_val_rec(value[assignment[variables[0]]],
                                     variables[1:], assignment)

Prob is a factor that represents a conditional probability by enumerating all of the values.

class Prob(CPD, TabFactor):
    """A factor defined by a conditional probability table""
    def __init__(self, var, pars, cpt):
        """Creates a factor from a conditional probability table, cpt
        The cpt values are assumed to be for the ordering par+[var]
        ""
        TabFactor.__init__(self, pars+[var], cpt)
        self.child = var
        self.parents = pars

9.4 Graphical Models

A graphical model consists of a set of variables and a set of factors. A belief network is a graphical model where all of the factors represent conditional probabilities. There are some operations (such as pruning variables) which are applicable to belief networks, but are not applicable to more general models. At the moment, we will treat them as the same.

from display import Displayable
from probFactors import CPD
import matplotlib.pyplot as plt

class GraphicalModel(Displayable):
    """The class of graphical models.
    A graphical model consists of a title, a set of variables and a set of factors."""
A belief network (also known as a Bayesian network) is a graphical model where all of the factors are conditional probabilities, and every variable has a conditional probability of it given its parents. This only checks the first condition, and builds some useful data structures.

The following creates a topological sort of the nodes, where the parents of a node come before the node in the resulting order. This is based on Kahn's algorithm from 1962.
The show method uses matplotlib to show the graphical structure of a belief network.

```python
def show(self):
    plt.ion()  # interactive
    ax = plt.figure().gca()
    ax.set_axis_off()
    plt.title(self.title)
    bbox = dict(boxstyle="round4,pad=1.0,rounding_size=0.5")
    for var in reversed(self.topological_sort()):
        if self.var2parents[var]:
            for par in self.var2parents[var]:
                ax.annotate(var.name, par.position, xytext=var.position,
                arrowprops={"arrowstyle": '<-"'},bbox=bbox,
                ha='center')
        else:
            x,y = var.position
            plt.text(x,y,var.name,bbox=bbox,ha='center')
```

9.4.1 Example Belief Networks

A Chain of 4 Variables

The first example belief network is a simple chain $A \rightarrow B \rightarrow C \rightarrow D$.

Please do not change this, as it is the example used for testing.

```python
from probVariables import Variable
from probFactors import Prob, LogisticRegression, NoisyOR

boolean = [False, True]
A = Variable("A", boolean, position=(0,0.8))
B = Variable("B", boolean, position=(0.333,0.6))
C = Variable("C", boolean, position=(0.666,0.4))
D = Variable("D", boolean, position=(1,0.2))
f_a = Prob(A,[],[0.4,0.6])
```
9.4. Graphical Models

Figure 9.1: The report-of-leaving belief network

```
91 f_b = Prob(B,[A],[[0.9,0.1],[0.2,0.8]])
92 f_c = Prob(C,[B],[[0.6,0.4],[0.3,0.7]])
93 f_d = Prob(D,[C],[[0.1,0.9],[0.75,0.25]])
94 bn_4ch = BeliefNetwork("4-chain", {A,B,C,D}, {f_a,f_b,f_c,f_d})
```

Report-of-Leaving Example

The second belief network, bn_report, is Example 8.15 of Poole and Mackworth [2017] (http://artint.info). The output of bn_report.show() is shown in Figure 9.1 of this document.

```
97 # Belief network report-of-leaving example (Example 8.15 shown in Figure 8.3) of
98 # Poole and Mackworth, Artificial Intelligence, 2017 http://artint.info
99 
100 Alarm = Variable("Alarm", boolean, position=(0.366,0.633))
101 Fire = Variable("Fire", boolean, position=(0.633,0.9))
102 Leaving = Variable("Leaving", boolean, position=(0.366,0.366))
103 Report = Variable("Report", boolean, position=(0.366,0.1))
104 Smoke = Variable("Smoke", boolean, position=(0.9,0.633))
105 Tamper = Variable("Tamper", boolean, position=(0.1,0.9))
```

http://aipython.org   Version 0.9.4    March 2, 2022
Sprinkler Example

The third belief network is the sprinkler example from Pearl. The output of `bn_sprinkler.show()` is shown in Figure 9.2 of this document.

```python
Season = Variable("Season", ["summer","winter"], position=(0.5,0.9))
Sprinkler = Variable("Sprinkler", ["on","off"], position=(0.9,0.6))
Rained = Variable("Rained", boolean, position=(0.1,0.6))
Grass_wet = Variable("Grass wet", boolean, position=(0.5,0.3))
Grass_shiny = Variable("Grass shiny", boolean, position=(0.1,0))
Shoes_wet = Variable("Shoes wet", boolean, position=(0.9,0))
```

http://aipython.org

Version 0.9.4

March 2, 2022
9.4. Graphical Models

Bipartite Diagnostic Model with Noisy-or

The belief network `bn_no1` is a bipartite diagnostic model, with independent diseases, and the symptoms depend on the diseases, where the CPDs are defined using noisy-or. Bipartite means it is in two parts; the diseases are only connected to the symptoms and the symptoms are only connected to the diseases. The output of `bn_no1.show()` is shown in Figure 9.3 of this document.

---

http://aipython.org  Version 0.9.4  March 2, 2022
Bipartite Diagnostic Model with Logistic Regression

The belief network `bn_lr1` is a bipartite diagnostic model, with independent diseases, and the symptoms depend on the diseases, where the CPDs are defined using logistic regression. It has the same graphical structure as the previous example (see Figure 9.3). This has the (approximately) the same conditional probabilities as the previous example when zero or one diseases are present. Note that $\text{sigmoid}(-2.2) \approx 0.1$
9.5. Inference Methods

Each of the inference methods implements the query method that computes the posterior probability of a variable given a dictionary of \{variable : value\} observations. The methods are Displayable because they implement the display method which is currently text-based.

```python
from display import Displayable
class InferenceMethod(Displayable):
    """The abstract class of graphical model inference methods""
    method_name = "unnamed" # each method should have a method name
    
    def __init__(self, gm=None):
        self.gm = gm

    def query(self, qvar, obs={}):
        """returns a \{value:prob\} dictionary for the query variable""
```

http://aipython.org  Version 0.9.4  March 2, 2022
We use bn_4ch as the test case, in particular \( P(B \mid D = \text{true}) \). This needs an error threshold, particularly for the approximate methods, where the default threshold is much too accurate.

### 9.6 Recursive Conditioning

An instance of a RC object takes in a graphical model. The query method uses recursive conditioning to compute the probability of a query variable given observations on other variables.
else:
    if split_order == None:
        split_order = [v for v in self.gm.variables if (v not in obs) and v != qvar]
    unnorm = [self.prob_search(dict_union({qvar:val},obs),
        self.gm.factors, split_order)
        for val in qvar.domain]
    p_obs = sum(unnorm)
    return {val:pr/p_obs for val,pr in zip(qvar.domain, unnorm)}

The following is the naive search-based algorithm. It is exponential in the number of variables, so is not very useful. However, it is simple, and useful to understand before looking at the more complicated algorithm used in the subclass.

```python
def prob_search(self, context, factors, split_order):
    """simple search algorithm
    context is a variable:value dictionary
    factors is a set of factors
    split_order is a list of variables in factors not assigned in context
    returns sum over variable assignments to variables in split order
    or product of factors """
    self.display(2,"calling prob_search",(context,factors))
    if not factors:
        return 1
    elif to_eval := {fac for fac in factors if fac.can_evaluate(context)}:
        # evaluate factors when all variables are assigned
        self.display(3,"prob_search evaluating factors",to_eval)
        val = math.prod(fac.get_value(context) for fac in to_eval)
        return val * self.prob_search(context, factors-to_eval, split_order)
    else:
        total = 0
        var = split_order[0]
        self.display(3, "prob_search branching on", var)
        for val in var.domain:
            total += self.prob_search(dict_union((var:val),context),
                factors, split_order[1:])
        self.display(3, "prob_search branching on", var,"returning", total)
        return total
```

The **recursive conditioning** algorithm adds forgetting and caching and recognizing disconnected components. We do this by adding a cache and redefining the recursive search algorithm. It inherits the query method.
def __init__(self, gm=None):
    self.cache = {(frozenset(), frozenset()):1}
    ProbSearch.__init__(self, gm)

def prob_search(self, context, factors, split_order):
    """ returns the number \sum_{split_order} \prod_{factors} given 
    assignments in context 
    context is a variable:value dictionary 
    factors is a set of factors 
    split_order is a list of variables in factors that are not assigned 
    in context 
    returns sum over variable assignments to variables in split_order 
    of the product of factors 
    """
    self.display(3, "calling rc", (context, factors))
    ce = (frozenset(context.items()), frozenset(factors)) # key for the 
    cache entry
    if ce in self.cache:
        self.display(3, "rc cache lookup", (context, factors))
        return self.cache[ce]
    # if not factors: # no factors; needed if you don't have forgetting 
    # and caching
    # return 1
    elif vars_not_in_factors := {var for var in context 
        if not any(var in fac.variables for 
        fac in factors)}:
        # forget variables not in any factor
        self.display(3, "rc forgetting variables", vars_not_in_factors)
        return self.prob_search({key:val for (key,val) in 
            context.items()
            if key not in vars_not_in_factors}, 
            factors, split_order)
    elif to_eval := {fac for fac in factors if 
        fac.can_evaluate(context)}:
        # evaluate factors when all variables are assigned
        self.display(3, "rc evaluating factors", to_eval)
        val = math.prod(fac.get_value(context) for fac in to_eval)
        if val == 0:
            return 0
        else:
            return val * self.prob_search(context, {fac for fac in factors 
                if fac not in to_eval}, 
                split_order)
    elif len(comp := connected_components(context, factors, 
        split_order)) > 1:
        # there are disconnected components
        self.display(3, "splitting into connected components", comp, "in 
        context", context)
        return math.prod(self.prob_search(context, f, eo) for (f, eo) in 
            comp)}

http://aipython.org
Version 0.9.4 March 2, 2022
else:
    assert split_order, "split_order should not be empty to get here"
    total = 0
    var = split_order[0]
    self.display(3, "rc branching on", var)
    for val in var.domain:
        total += self.prob_search(dict_union({var:val},context),
                factors, split_order[1:])
    self.cache[ce] = total
    self.display(2, "rc branching on", var,"returning", total)
    return total

connected_components returns a list of connected components, where a connected component is a set of factors and a set of variables, where the graph that connects variables and factors that involve them is connected. The connected components are built one at a time; with a current connected component. At all times factors is partitioned into 3 disjoint sets:

- component_factors containing factors in the current connected component where all factors that share a variable are already in the component

- factors_to_check containing factors in the current connected component where potentially some factors that share a variable are not in the component; these need to be checked

- other_factors the other factors that are not (yet) in the connected component

---

```python
def connected_components(context, factors, split_order):
    """returns a list of (f,e) where f is a subset of factors and e is a subset of split_order such that each element shares the same variables that are disjoint from other elements.
    """
    other_factors = set(factors) #copies factors
    factors_to_check = {other_factors.pop()} # factors in connected component still to be checked
    component_factors = set() # factors in first connected component already checked
    component_variables = set() # variables in first connected component
    while factors_to_check:
        next_fac = factors_to_check.pop()
        component_factors.add(next_fac)
        new_vars = set(next_fac.variables) - component_variables - context.keys()
        component_variables |= new_vars
        for var in new_vars:
            ...
factors_to_check |= \{f \text{ for } f \text{ in } \text{other_factors} \text{ if } \text{var} \text{ in } f.\text{variables}\}
other_factors -= factors_to_check # set difference
if other_factors:
    return ([(component_factors, [e \text{ for } e \text{ in } \text{split_order} \text{ if } e \text{ in } component_variables])
        + \text{connected_components(context, other_factors, [e \text{ for } e \text{ in } split_order} 
        \text{ if } e \not\text{ in } \text{component_variables}])
    )
else:
    return [(component_factors, \text{split_order}]]

Testing:

```python

from probGraphicalModels import bn_4ch, A,B,C,D,f_a,f_b,f_c,f_d

bn_4chv = ProbRC(bn_4ch)
## bn_4chv.query(A,{})
## bn_4chv.query(D,{})
## InferenceMethod.max_display_level = 3 # show more detail in displaying
## InferenceMethod.max_display_level = 1 # show less detail in displaying
## bn_4chv.query(A,{D:True},[C,B])
## bn_4chv.query(B,{A:True,D:False})

from probGraphicalModels import bn_report, Alarm, Fire, Leaving, Report, Smoke, Tamper

bn_reportRC = ProbRC(bn_report) # answers queries using recursive conditioning
## bn_reportRC.query(Tamper,{}),
## InferenceMethod.max_display_level = 0 # show no detail in displaying
## bn_reportRC.query(Leaving,{}),
## bn_reportRC.query(Tamper,{}, split_order=[Smoke,Fire,Alarm,Leaving,Report])
## bn_reportRC.query(Tamper,{Report:True})
## bn_reportRC.query(Tamper,{Report:True,Smoke:False})
## Note what happens to the cache when these are called in turn:
## bn_reportRC.query(Tamper,{Report:True}, split_order=[Smoke,Fire,Alarm,Leaving])
## bn_reportRC.query(Report:{}, split_order=[Tamper,Fire,Alarm,Leaving])

from probGraphicalModels import bn_sprinkler, Season, Sprinkler, Rained, Grass_wet, Grass_shiny, Shoes_wet

bn_sprinklerv = ProbRC(bn_sprinkler)
## bn_sprinklerv.query(Shoes_wet,{},)
## bn_sprinklerv.query(Shoes_wet,{Rained:True})
## bn_sprinklerv.query(Shoes_wet,{Grass_shiny:True})
## bn_sprinklerv.query(Shoes_wet,{Grass_shiny:False,Rained:True})
```
from probGraphicalModels import bn_no1, bn_lr1, Cough, Fever, Sneeze, Cold, Flu, Covid

bn_no1v = ProbRC(bn_no1)
bn_lr1v = ProbRC(bn_lr1)

## bn_no1v.query(Flu, {Fever:1, Sneeze:1})
## bn_lr1v.query(Flu, {Fever:1, Sneeze:1})
## bn_lr1v.query(Cough,{})
## bn_lr1v.query(Cold,{Cough:1,Sneeze:0,Fever:1})
## bn_lr1v.query(Flu,{Cough:0,Sneeze:1,Fever:1})
## bn_lr1v.query(Covid,{Cough:1,Sneeze:0,Fever:1})
## bn_lr1v.query(Covid,{Cough:1,Sneeze:0,Fever:1,Flu:0})
## bn_lr1v.query(Covid,{Cough:1,Sneeze:0,Fever:1,Flu:1})

if __name__ == "__main__":
    InferenceMethod.testIM(ProbRC)

9.7 Variable Elimination

An instance of a VE object takes in a graphical model. The query method uses variable elimination to compute the probability of a variable given observations on some other variables.

```python
class VE(InferenceMethod):
    """The class that queries Graphical Models using variable elimination.
    gm is graphical model to query
    """
    method_name = "variable elimination"

    def __init__(self,gm=None):
        InferenceMethod.__init__(self, gm)

    def query(self,var,obs={},elim_order=None):
        """computes P(var|obs) where var is a variable obs is a {variable:value} dictionary""
        if var in obs:
            return {var:1 if val == obs[var] else 0 for val in var.domain}
        else:
            if elim_order == None:
                elim_order = self.gm.variables
            projFactors = [self.project_observations(fact,obs)
                           for fact in self.gm.factors]
            for v in elim_order:
                if v != var and v not in obs:
```

http://aipython.org  Version 0.9.4  March 2, 2022
9. Reasoning Under Uncertainty

```python
projFactors = self.eliminate_var(projFactors,v)
unnorm = factor_times(var,projFactors)
p_obs = sum(unnorm)
self.display(1,"Unnormalized probs:",unnorm,"Prob obs:",p_obs)
return {val:pr/p_obs for val,pr in zip(var.domain, unnorm)}
```

A FactorObserved is a factor that is the result of some observations on another factor. We don’t store the values in a list; we just look them up as needed. The observations can include variables that are not in the list, but should have some intersection with the variables in the factor.

```python
class FactorObserved(Factor):
    def __init__(self,factor,obs):
        Factor.__init__(self, [v for v in factor.variables if v not in obs])
        self.observed = obs
        self.orig_factor = factor

    def get_value(self,assignment):
        ass = assignment.copy()
        for ob in self.observed:
            ass[ob]=self.observed[ob]
        return self.orig_factor.get_value(ass)
```

A FactorSum is a factor that is the result of summing out a variable from the product of other factors. I.e., it constructs a representation of:

\[
\sum \prod_{f \in \text{factors}} f.
\]

We store the values in a list in a lazy manner; if they are already computed, we used the stored values. If they are not already computed we can compute and store them.

```python
class FactorSum(Factor):
    def __init__(self,var,factors):
        self.var_summed_out = var
        self.factors = factors
        vars = []
        for fac in factors:
            for v in fac.variables:
                if v is not var and v not in vars:
                    vars.append(v)
        Factor.__init__(self,vars)
        self.values = {}

    def get_value(self,assignment):
        """lazy implementation: if not saved, compute it. Return saved value""
        asst = frozenset(assignment.items())
```

http://aipython.org
9.7. Variable Elimination

```python
if asst in self.values:
    return self.values[asst]
else:
    total = 0
    new_asst = assignment.copy()
    for val in self.var_summed_out.domain:
        new_asst[self.var_summed_out] = val
        total += math.prod(fac.get_value(new_asst) for fac in self.factors)
    self.values[asst] = total
    return total
```

The method `factor_times` multiples a set of factors that are all factors on the same variable (or on no variables). This is the last step in variable elimination before normalizing. It returns an array giving the product for each value of `variable`.

```python
def factor_times(variable, factors):
    """when factors are factors just on variable (or on no variables)""
    prods = []
    facs = [f for f in factors if variable in f.variables]
    for val in variable.domain:
        ast = {variable:val}
        prods.append(math.prod(f.get_value(ast) for f in facs))
    return prods
```

To project observations onto a factor, for each variable that is observed in the factor, we construct a new factor that is the factor projected onto that variable. `Factor_observed` creates a new factor that is the result is assigning a value to a single variable.

```python
def project_observations(self,factor,obs):
    """Returns the resulting factor after observing obs"
    obs is a dictionary of {variable:value} pairs.
    """
    if any((var in obs) for var in factor.variables):
        # a variable in factor is observed
        return FactorObserved(factor,obs)
    else:
        return factor
```

```python
def eliminate_var(self,factors,var):
    """Eliminate a variable var from a list of factors.
    Returns a new set of factors that has var summed out.
    """
    self.display(2,"eliminating ",str(var))
    contains_var = []
    not_contains_var = []
    for fac in factors:
```

http://aipython.org
if var in fac.variables:
    contains_var.append(fac)
else:
    not_contains_var.append(fac)
if contains_var == []:
    return factors
else:
    newFactor = FactorSum(var,contains_var)
    self.display(2, "Multiplying:"
str(f) for f in contains_var])
    self.display(2, "Creating factor: ", newFactor)
    self.display(3, newFactor.to_table() # factor in detail
    not_contains_var.append(newFactor)
return not_contains_var

from probGraphicalModels import bn_4ch, A,B,C,D
bn_4chv = VE(bn_4ch)
## bn_4chv.query(A,{})
## bn_4chv.query(D,{})
## InferenceMethod.max_display_level = 3 # show more detail in displaying
## InferenceMethod.max_display_level = 1 # show less detail in displaying
## bn_4chv.query(A,{D:True})
## bn_4chv.query(B,{A:True,D:False})

from probGraphicalModels import
bn_reportv = VE(bn_report) # answers queries using variable elimination
## bn_reportv.query(Tamper,{})
## InferenceMethod.max_display_level = 0 # show no detail in displaying
## bn_reportv.query(Leaving,{},elim_order=[Smoke,Report,Leaving,Alarm,Fire])
## bn_reportv.query(Tamper,{Report:True})
## bn_reportv.query(Tamper,{Report:True,Smoke:False})

from probGraphicalModels import bn_sprinkler, Season, Sprinkler, Rained,
Grass_wet, Grass_shiny, Shoes_wet
bn_sprinklerv = VE(bn_sprinkler)
## bn_sprinklerv.query(Shoes_wet,{})
## bn_sprinklerv.query(Shoes_wet,{Rained:True})
## bn_sprinklerv.query(Shoes_wet,{Grass_shiny:True})
## bn_sprinklerv.query(Shoes_wet,{Grass_shiny:False,Rained:True})

from probGraphicalModels import bn_lr1, Cough, Fever, Sneeze, Cold, Flu,
Covid
vediag = VE(bn_lr1)
## vediag.query(Cough,{},elim_order=[Cough,Flu,Sneeze,Cold,Fever])
## vediag.query(Leaving,Cold)
9.8 Stochastic Simulation

9.8.1 Sampling from a discrete distribution

The method `sample_one` generates a single sample from a (possible unnormalized) distribution. `dist` is a `{value : weight}` dictionary, where `weight ≥ 0`. This returns a value with probability in proportion to its weight.

```python
import random
from probGraphicalModels import InferenceMethod

def sample_one(dist):
    
    """returns the index of a single sample from normalized distribution dist.""
    
    rand = random.random() * sum(dist.values())
    cum = 0  # cumulative weights
    for v in dist:
        cum += dist[v]
        if cum > rand:
            return v
```

If we want to generate multiple samples, repeatedly calling `sample_one` may not be efficient. If we want to generate `n` samples, and the distribution is over `m` values, `sample_one` takes time $O(mn)$. If $m$ and $n$ are of the same order of magnitude, we can do better.

The method `sample_multiple` generates multiple samples from a distribution defined by `dist`, where `dist` is a `{value : weight}` dictionary, where `weight ≥ 0` and the weights cannot all be zero. This returns a list of values, of length `num_samples`, where each sample is selected with a probability proportional to its weight.

The method generates all of the random numbers, sorts them, and then goes through the distribution once, saving the selected samples.

```python
def sample_multiple(dist, num_samples):
    
    """returns a list of num_samples values selected using distribution dist.
    dist is a {value:weight} dictionary that does not need to be normalized
    ""
    total = sum(dist.values())
    rands = sorted(random.random()*total for i in range(num_samples))
    result = []
    dist_items = list(dist.items())
```
Exercise 9.1

What is the time and space complexity the following 4 methods to generate \( n \) samples, where \( m \) is the length of \( dist \):

(a) \( n \) calls to \texttt{sample_one}

(b) \texttt{sample_multiple}

(c) Create the cumulative distribution (choose how this is represented) and, for each random number, do a binary search to determine the sample associated with the random number.

(d) Choose a random number in the range \( [i/n, (i+1)/n) \) for each \( i \in \text{range}(n) \), where \( n \) is the number of samples. Use these as the random numbers to select the particles. (Does this give random samples?)

For each method suggest when it might be the best method.

The \texttt{test_sampling} method can be used to generate the statistics from a number of samples. It is useful to see the variability as a function of the number of samples. Try it for few samples and also for many samples.

```python
# try the following queries a number of times each:
# test_sampling({1:1,2:2,3:3,4:4}, 100)
# test_sampling({1:1,2:2,3:3,4:4}, 100000)
```

### 9.8.2 Sampling Methods for Belief Network Inference

A \texttt{SamplingInferenceMethod} is an \texttt{InferenceMethod}, but the query method also takes arguments for the number of samples and the sample-order (which is an ordering of factors). The first methods assume a belief network (and not an undirected graphical model).
9.8. Stochastic Simulation

```python
class SamplingInferenceMethod(InferenceMethod):
    """The abstract class of sampling-based belief network inference methods""

def __init__(self, gm=None):
    InferenceMethod.__init__(self, gm)

def query(self, qvar, obs={}, number_samples=1000, sample_order=None):
    raise NotImplementedError("SamplingInferenceMethod query") # abstract

9.8.3 Rejection Sampling

```
if not rejected:
    counts[sample[qvar]] += 1
    self.display(2,"Accepted")

tot = sum(counts.values())
# As well as the distribution we also include raw counts
dist = {c:v/tot if tot>0 else 1/len(qvar.domain) for (c,v) in counts.items()}
dist["raw_counts"] = counts
return dist

9.8.4 Likelihood Weighting

Likelihood weighting includes a weight for each sample. Instead of rejecting samples based on observations, likelihood weighting changes the weights of the sample in proportion with the probability of the observation. The weight then becomes the probability that the variable would have been rejected.

```python
class LikelihoodWeighting(SamplingInferenceMethod):
    """The class that queries Graphical Models using Likelihood weighting."

    gm is a belief network to query
    ""
    method_name = "likelihood weighting"

    def __init__(self, gm=None):
        SamplingInferenceMethod.__init__(self, gm)

    def query(self,qvar,obs={},number_samples=1000,sample_order=None):
        """computes P(qvar | obs) where
        qvar is a variable.
        obs is a {variable:value} dictionary.
        sample_order is a list of factors where factors defining the parents
        come before the factors for the child.
        ""
        if sample_order is None:
            sample_order = self.gm.topological_sort()
            self.display(2,*[v for v in sample_order
                if v not in obs],sep="\t")
            counts = {val:0 for val in qvar.domain}
        for i in range(number_samples):
            sample = {}
            weight = 1.0
            for nvar in sample_order:
                fac = self.gm.var2cpt[nvar]
                if nvar in obs:
                    sample[nvar] = obs[nvar]
                    weight *= fac.get_value(sample)
                else:
                    val = sample_one({v:fac.get_value(**sample,nvar:v)) for v in nvar.domain})
```

[http://aipython.org](http://aipython.org)
9.8. Stochastic Simulation

```python
self.display(2,val,end="\t")
sample[nvar] = val
counts[sample[qvar]] += weight
self.display(2,weight)
tot = sum(counts.values())
# as well as the distribution we also include the raw counts
dist = {c:v/tot for (c,v) in counts.items()}
dist["raw_counts"] = counts
return dist
```

**Exercise 9.2** Change this algorithm so that it does importance sampling using a proposal distribution. It needs `sample_one` using a different distribution and then update the weight of the current sample. For testing, use a proposal distribution that only specifies probabilities for some of the variables (and the algorithm uses the probabilities for the network in other cases).

9.8.5 Particle Filtering

In this implementation, a particle is a `{variable : value}` dictionary. Because adding a new value to dictionary involves a side effect, the dictionaries need to be copied during resampling.

```python
class ParticleFiltering(SamplingInferenceMethod):
    """The class that queries Graphical Models using Particle Filtering.
    gm is a belief network to query
    """
    method_name = "particle filtering"

    def __init__(self, gm=None):
        SamplingInferenceMethod.__init__(self, gm)

    def query(self, qvar, obs={}, number_samples=1000, sample_order=None):
        """computes P(qvar | obs) where
        qvar is a variable.
        obs is a {variable:value} dictionary.
        sample_order is a list of factors where factors defining the parents
        come before the factors for the child.
        """
        if sample_order is None:
            sample_order = self.gm.topological_sort()
        self.display(2,*[v for v in sample_order
        if v not in obs],sep="\t")
        particles = [{()} for i in range(number_samples)]
        for nvar in sample_order:
            fac = self.gm.var2cpt[nvar]
            if nvar in obs:
                weights = [fac.get_value(**part, nvar=obs[nvar]) for part
                in particles]
```

http://aipython.org  Version 0.9.4  March 2, 2022
particles = [[**p, nvar:obs[nvar]] for p in resample(particles, weights, number_samples)]
else:
    for part in particles:
        part[nvar] = sample_one({v:fac.get_value(**part, nvar:v) for v in nvar.domain})
        self.display(2,part[nvar],end="\t")
counts = {val:0 for val in qvar.domain}
for part in particles:
    counts[part[qvar]] += 1
tot = sum(counts.values())
# as well as the distribution we also include the raw counts
dist = {c:v/tot for (c,v) in counts.items()}
dist["raw_counts"] = counts
return dist

Resampling

Resample is based on sample_multiple but works with an array of particles.
(Aside: Python doesn’t let us use sample_multiple directly as it uses a dictionary, and particles, represented as dictionaries can’t be the key of dictionaries).

```python
def resample(particles, weights, num_samples):
    """returns num_samples copies of particles resampled according to
    weights.
    particles is a list of particles
    weights is a list of positive numbers, of same length as particles
    num_samples is an integer
    """
    total = sum(weights)
rands = sorted(random.random()*total for i in range(num_samples))
result = []
cum = weights[0]  # cumulative sum
index = 0
for r in rands:
    while r>cum:
        index += 1
        cum += weights[index]
    result.append(particles[index])
return result
```

9.8.6 Examples

```python
from probGraphicalModels import bn_4ch, A,B,C,D
bn_4chr = RejectionSampling(bn_4ch)
bn_4chL = LikelihoodWeighting(bn_4ch)
```

http://aipython.org  Version 0.9.4  March 2, 2022
9.8. Stochastic Simulation

```python
## InferenceMethod.max_display_level = 2 # detailed tracing for all inference methods
## bn_4chr.query(A,{})
## bn_4chr.query(C,{})
## bn_4chr.query(A,{C:True})
## bn_4chr.query(B,{A:True,C:False})

from probGraphicalModels import bn_report, Alarm, Fire, Leaving, Report, Smoke, Tamper

bn_reportr = RejectionSampling(bn_report) # answers queries using rejection sampling
bn_reportL = LikelihoodWeighting(bn_report) # answers queries using likelihood weighting
bn_reportp = ParticleFiltering(bn_report) # answers queries using particle filtering

## bn_reportr.query(Tamper,{},)
## bn_reportr.query(Tamper,{},)
## bn_reportr.query(Tamper,{Report:True})
## InferenceMethod.max_display_level = 0 # no detailed tracing for all inference methods
## bn_reportr.query(Tamper,{Report:True},number_samples=100000)
## bn_reportr.query(Tamper,{Report:True,Smoke:False})
## bn_reportr.query(Tamper,{Report:True,Smoke:False},number_samples=100)

## bn_reportL.query(Tamper,{Report:True,Smoke:False},number_samples=100)
## bn_reportL.query(Tamper,{Report:True,Smoke:False},number_samples=100)

from probGraphicalModels import bn_sprinkler, Season, Sprinkler
from probGraphicalModels import Rained, Grass_wet, Grass_shiny, Shoes_wet

bn_sprinklerr = RejectionSampling(bn_sprinkler) # answers queries using rejection sampling
bn_sprinklerL = LikelihoodWeighting(bn_sprinkler) # answers queries using rejection sampling
bn_sprinklerp = ParticleFiltering(bn_sprinkler) # answers queries using particle filtering

#bn_sprinklerr.query(Shoes_wet,{Grass_shiny:True,Rained:True})
#bn_sprinklerL.query(Shoes_wet,{Grass_shiny:True,Rained:True})
#bn_sprinklerp.query(Shoes_wet,{Grass_shiny:True,Rained:True})

if __name__ == "__main__":
    InferenceMethod.testIM(RejectionSampling, threshold=0.1)
    InferenceMethod.testIM(LikelihoodWeighting, threshold=0.1)
    InferenceMethod.testIM(ParticleFiltering, threshold=0.1)
```

**Exercise 9.3** This code keeps regenerating the distribution of a variable given its parents. Implement one or both of the following, and compare them to the original. Make `cond_dist` return a slice that corresponds to the distribution, and then use the slice instead of the dictionary (a list slice does not generate new data structures). Make `cond_dist` remember values it has already computed, and only return these.
9.8.7 Gibbs Sampling

The following implements **Gibbs sampling**, a form of **Markov Chain Monte Carlo** MCMC.

```python
# import random
from probGraphicalModels import InferenceMethod
from probStochSim import sample_one, SamplingInferenceMethod

class GibbsSampling(SamplingInferenceMethod):
    """The class that queries Graphical Models using Gibbs Sampling.
    """
    bn is a graphical model (e.g., a belief network) to query
    """
    method_name = "Gibbs sampling"

    def __init__(self, gm=None):
        SamplingInferenceMethod.__init__(self, gm)
        self.gm = gm

    def query(self, qvar, obs={}, number_samples=1000, burn_in=100,
        sample_order=None):
        """computes P(qvar | obs) where
        qvar is a variable.
        obs is a {variable:value} dictionary.
        sample_order is a list of non-observed variables in order, or
        if sample_order None, the variables are shuffled at each iteration.
        """
        counts = {val:0 for val in qvar.domain}
        if sample_order is not None:
            variables = sample_order
        else:
            variables = [v for v in self.gm.variables if v not in obs]
        var_to_factors = {v:set() for v in self.gm.variables}
        for fac in self.gm.factors:
            for var in fac.variables:
                var_to_factors[var].add(fac)
        sample = {var:random.choice(var.domain) for var in variables}
        self.display(2,"Sample: ",sample)
        sample.update(obs)
        for i in range(burn_in + number_samples):  
            if sample_order == None:
                random.shuffle(variables)
            for var in variables:
                # get unnormalized probability distribution of var given its
                # neighbours
                vardist = {val:1 for val in var.domain}
            for val in var.domain:
                sample[var] = val
```

http://aipython.org
for fac in var_to_factors[var]: # Markov blanket
    vardist[val] *= fac.get_value(sample)
    sample[var] = sample_one(vardist)
if i >= burn_in:
    counts[sample[qvar]] +=1
    tot = sum(counts.values())
# as well as the computed distribution, we also include raw counts
    dist = {c:v/tot for (c,v) in counts.items()}
    dist["raw_counts"] = counts
return dist

#from probGraphicalModels import bn_4ch, A,B,C,D
bn_4chg = GibbsSampling(bn_4ch)
## InferenceMethod.max_display_level = 2 # detailed tracing for all inference methods
bn_4chg.query(A,{})
## bn_4chg.query(D,{})
## bn_4chg.query(B,{D:True})
## bn_4chg.query(B,{A:True,C:False})

from probGraphicalModels import bn_report, Alarm, Fire, Leaving, Report, Smoke, Tamper
bn_reportg = GibbsSampling(bn_report)
## bn_reportg.query(Tamper,{Report:True},number_samples=1000)

if __name__ == "__main__":
    InferenceMethod.testIM(GibbsSampling, threshold=0.1)

Exercise 9.4 Change the code so that it can have multiple query variables. Make
the list of query variable be an input to the algorithm, so that the default value is
the list of all non-observed variables.

Exercise 9.5 In this algorithm, explain where it computes the probability of a
variable given its Markov blanket. Instead of returning the average of the samples
for the query variable, it is possible to return the average estimate of the probabil-
ity of the query variable given its Markov blanket. Does this converge to the same
answer as the given code? Does it converge faster, slower, or the same?

9.8.8 Plotting Behaviour of Stochastic Simulators

The stochastic simulation runs can give different answers each time they are run. For the algorithms that give the same answer in the limit as the number of samples approaches infinity (as do all of these algorithms), the algorithms can be compared by comparing the accuracy for multiple runs. Summary statistics like the variance may provide some information, but the assumptions behind the variance being appropriate (namely that the distribution is approximately Gaussian) may not hold for cases where the predictions are bounded and often skewed.

http://aipython.org

Version 0.9.4 March 2, 2022
9. Reasoning Under Uncertainty

It is more appropriate to plot the distribution of predictions over multiple runs. The `plot_stats` method plots the prediction of a particular variable (or for the partition function) for a number of runs of the same algorithm. On the x-axis, is the prediction of the algorithm. On the y-axis is the number of runs with prediction less than or equal to the x value. Thus this is like a cumulative distribution over the predictions, but with counts on the y-axis.

Note that for runs where there are no samples that are consistent with the observations (as can happen with rejection sampling), the prediction of probability is 1.0 (as a convention for 0/0).

That variable `what` contains the query variable, or `what` is "`prob.ev`", the probability of evidence.

---

```python
import matplotlib.pyplot as plt

def plot_stats(method, qvar, qval, obs, number_runs=1000, **queryargs):
    """Plots a cumulative distribution of the prediction of the model.
    method is a InferenceMethod (that implements appropriate query(.))
    plots P(qvar=qval | obs)
    qvar is the query variable, qval is corresponding value
    obs is the {variable:value} dictionary representing the observations
    number_iterations is the number of runs that are plotted
    **queryargs is the arguments to query (often number_samples for
    sampling methods)
    """
    plt.ion()
    plt.xlabel("value")
    plt.ylabel("Cumulative Number")
    method.max_display_level, prev_mdl = 0, method.max_display_level #no display
    answers = [method.query(qvar,obs,**queryargs)
                for i in range(number_runs)]
    values = [ans[qval] for ans in answers]
    label = f"{method.method_name} P({qvar}={qval}|{', '.join(f'{var}={val}'
                 for (var,val) in obs.items())})"
    values.sort()
    plt.plot(values,range(number_runs),label=label)
    plt.legend() #loc="upper left")
    plt.draw()
    method.max_display_level = prev_mdl # restore display level

# Try:
# plot_stats(bn_reportr,Tamper,True,{Report:True,Smoke:True},number_samples=1000,
# number_runs=1000)
#
# plot_stats(bn_reportr,Tamper,True,{Report:True,Smoke:True},number_samples=1000,
# number_runs=1000)
```
9.9. Hidden Markov Models

This code for hidden Markov models is independent of the graphical models code, to keep it simple. Section 9.10 gives code that models hidden Markov models, and more generally, dynamic belief networks, using the graphical models code.

This HMM code assumes there are multiple Boolean observation variables that depend on the current state and are independent of each other given the state.

```
plot_stats(bn_reportp,Tamper,True,{Report:True,Smoke:True},number_samples=1000,
number_runs=1000)
#
plot_stats(bn_reportr,Tamper,True,{Report:True,Smoke:True},number_samples=100,
number_runs=1000)
#
plot_stats(bn_reportL,Tamper,True,{Report:True,Smoke:True},number_samples=100,
number_runs=1000)
#
plot_stats(bn_reportg,Tamper,True,{Report:True,Smoke:True},number_samples=1000,
number_runs=1000)

def plot_mult(methods, example, qvar, qval, obs, number_samples=1000,
number_runs=1000):
    for method in methods:
        solver = method(example)
        if isinstance(method,SamplingInferenceMethod):
            plot_stats(solver, qvar, qval, obs, number_samples, number_runs)
        else:
            plot_stats(solver, qvar, qval, obs, number_runs)

from probRC import ProbRC
# Try following (but it takes a while..)
methods =
    [ProbRC,RejectionSampling,LikelihoodWeighting,ParticleFiltering,GibbsSampling]
#plot_mult(methods,bn_report,Tamper,True,{Report:True,Smoke:False},number_samples=100,
#number_runs=1000)
#
plot_mult(methods,bn_report,Tamper,True,{Report:False,Smoke:True},number_samples=100,
number_runs=1000)
#
# Sprinkler Example:
#
plot_stats(bn_sprinklerr,Shoes_wet,True,{Grass_shiny:True,Rained:True},number_samples=1000)
#
plot_stats(bn_sprinklerL,Shoes_wet,True,{Grass_shiny:True,Rained:True},number_samples=1000)
```
import random
from probStochSim import sample_one, sample_multiple

class HMM(object):
    def __init__(self, states, obsvars, pobs, trans, indist):
        """A hidden Markov model.
        states - set of states
        obsvars - set of observation variables
        pobs - probability of observations, pobs[i][s] is P(Obs_i=True | State=s)
        trans - transition probability - trans[i][j] gives P(State=j | State=i)
        indist - initial distribution - indist[s] is P(State_0 = s)
        ""
        self.states = states
        self.obsvars = obsvars
        self.pobs = pobs
        self.trans = trans
        self.indist = indist

Consider the following example. Suppose you want to unobtrusively keep track of an animal in a triangular enclosure using sound. Suppose you have 3 microphones that provide unreliable (noisy) binary information at each time step. The animal is either close to one of the 3 points of the triangle or in the middle of the triangle.

# state
# 0=middle, 1,2,3 are corners
states1 = {'middle', 'c1', 'c2', 'c3'} # states
obs1 = {'m1', 'm2', 'm3'} # microphones

The observation model is as follows. If the animal is in a corner, it will be detected by the microphone at that corner with probability 0.6, and will be independently detected by each of the other microphones with a probability of 0.1. If the animal is in the middle, it will be detected by each microphone with a probability of 0.4.

# pobs gives the observation model:
#pobs[mi][state] is P(mi=on | state)
closeMic=0.6; farMic=0.1; midMic=0.4
pobs1 = {'m1': {'middle': midMic, 'c1': closeMic, 'c2': farMic, 'c3': farMic},
         # mic 1
         'm2': {'middle': midMic, 'c1': farMic, 'c2': closeMic, 'c3': farMic},
         # mic 2
         'm3': {'middle': midMic, 'c1': farMic, 'c2': farMic, 'c3': closeMic}} # mic 3

The transition model is as follows: If the animal is in a corner it stays in the same corner with probability 0.80, goes to the middle with probability 0.1
or goes to one of the other corners with probability 0.05 each. If it is in the middle, it stays in the middle with probability 0.7, otherwise it moves to one of the corners, each with probability 0.1.

# trans specifies the dynamics
# trans[i] is the distribution over states resulting from state i
# trans[i][j] gives P(S=j | S=i)
sm=0.7; mmc=0.1  # transition probabilities when in middle
sc=0.8; mcm=0.1; mcc=0.05  # transition probabilities when in a corner
trans1 = {'middle': { 'middle': sm, 'c1': mmc, 'c2': mmc, 'c3': mmc},  # was in middle
           'c1': { 'middle': mcm, 'c1': sc, 'c2': mcc, 'c3': mcc},  # was in corner
           'c2': { 'middle': mcm, 'c1': mcc, 'c2': sc, 'c3': mcc},  # was in corner
           'c3': { 'middle': mcm, 'c1': mcc, 'c2': mcc, 'c3': sc}}  # was in corner

Initially the animal is in one of the four states, with equal probability.

indist1 = {st:1.0/len(states1) for st in states1}

hmm1 = HMM(states1, obs1, pobs1, trans1, indist1)

9.9.1 Exact Filtering for HMMs

A HMMVEFilter has a current state distribution which can be updated by observing or by advancing to the next time.

class HMMVEFilter(Displayable):
    def __init__(self,hmm):
        self.hmm = hmm
        self.state_dist = hmm.indist
    def filter(self, obsseq):
        """updates and returns the state distribution following the sequence of observations in obsseq using variable elimination."

        Note that it first advances time.
        This is what is required if it is called sequentially.
        If that is not what is wanted initially, do an observe first.
        """
        for obs in obsseq:

http://aipython.org  Version 0.9.4  March 2, 2022
def observe(self, obs):
    """updates state conditioned on observations.
    obs is a list of values for each observation variable""
    for i in self.hmm.obsvars:
        self.state_dist = {st:self.state_dist[st]*(self.hmm.pobs[i][st]
                         if obs[i] else
                         (1-self.hmm.pobs[i][st]))
                         for st in self.hmm.states}

        norm = sum(self.state_dist.values()) # normalizing constant
        self.state_dist = {st:self.state_dist[st]/norm
                         for st in self.hmm.states}
    self.display(2,"After observing",obs,"state
distribution:",self.state_dist)

def advance(self):
    """advance to the next time""
    nextstate = {st:0.0
                         for st in self.hmm.states} # distribution over
    next states

        for j in self.hmm.states: # j ranges over next states
            for i in self.hmm.states: # i ranges over previous states
                nextstate[j] += self.hmm.trans[i][j]*self.state_dist[i]

        self.state_dist = nextstate

        self.display(2,"After advancing state
distribution:",self.state_dist)

The following are some queries for hmm1.

probHMM.py — (continued)
Exercise 9.6  The representation assumes that there are a list of Boolean observations. Extend the representation so that the each observation variable can have multiple discrete values. You need to choose a representation for the model, and change the algorithm.

9.9.2  Localization

The localization example in the book is a controlled HMM, where there is a given action at each time and the transition depends on the action. In this class, the transition is set to None initially, and needs to be provided with an action to determine the transition probability.

```python
from probHMM import HMMVEfilter, HMM
from display import Displayable
import matplotlib.pyplot as plt
from matplotlib.widgets import Button, CheckButtons

class HMM_Controlled(HMM):
    """A controlled HMM, where the transition probability depends on the action.
    Instead of the transition probability, it has a function act2trans from action to transition probability.
    Any algorithms need to select the transition probability according to the action.
    """
    def __init__(self, states, obsvars, pobs, act2trans, indist):
        self.act2trans = act2trans
        HMM.__init__(self, states, obsvars, pobs, None, indist)

local_states = list(range(16))
door_positions = {2,4,7,11}
def prob_door(loc): return 0.8 if loc in door_positions else 0.1
local_obs = {'door':[prob_door(i) for i in range(16)]}
act2trans = {'right': [[0.1 if next == current
                        else 0.8 if next == (current+1)%16
                        else 0.074 if next == (current+2)%16
                        else 0.002 for next in range(16)] for
                        current in range(16)],
            'left': [[0.1 if next == current
                      else 0.8 if next == (current-1)%16
                      else 0.074 if next == (current-2)%16
                      else 0.002 for next in range(16)] for
                      current in range(16)]}

http://aipython.org  Version 0.9.4  March 2, 2022
To change the VE localization code to allow for controlled HMMs, notice that the action selects which transition probability to use.

```python
class HMM_Local(HMMVEfilter):
    """VE filter for controlled HMMs """
    def __init__(self, hmm):
        HMMVEfilter.__init__(self, hmm)
        self.hmm.trans = self.hmm.act2trans[action]
        self.advance()

loc_filt = HMM_Local(hmm_16pos)
# loc_filt.observe({'door':True}); loc_filt.go("right");
# loc_filt.observe({'door':False}); loc_filt.go("right");
# loc_filt.observe({'door':True})
```

The following lets us interactively move the agent and provide observations. It shows the distribution over locations.

```python
class Show_Localization(Displayable):
    def __init__(self,hmm):
        self.hmm = hmm
        self.loc_filt = HMM_Local(hmm)
        fig,(self.ax) = plt.subplots()
        plt.subplots_adjust(bottom=0.2)
        left_butt = Button(plt.axes([0.05,0.02,0.1,0.05]), "left")
        left_butt.on_clicked(self.left)
        right_butt = Button(plt.axes([0.25,0.02,0.1,0.05]), "right")
        right_butt.on_clicked(self.right)
        door_butt = Button(plt.axes([0.45,0.02,0.1,0.05]), "door")
        door_butt.on_clicked(self.door)
        nodoor_butt = Button(plt.axes([0.65,0.02,0.1,0.05]), "no door")
        nodoor_butt.on_clicked(self.nodoor)
        reset_butt = Button(plt.axes([0.85,0.02,0.1,0.05]), "reset")
        reset_butt.on_clicked(self.reset)
        #this makes sure y-axis goes to 1, graph overwritten in draw_dist
        self.draw_dist()
        plt.show()
```

```python
def draw_dist(self):
    self.ax.clear()
    plt.ylim(0,1)
```
9.9. Hidden Markov Models

```python
self.ax.set_ylabel("Probability")
self.ax.set_xlabel("Location")
self.ax.set_title("Location Probability Distribution")
self.ax.set_xticks(self.hmm.states)
vals = [self.loc_filt.state_dist[i] for i in self.hmm.states]
self.bars = self.ax.bar(self.hmm.states, vals, color='black')
self.ax.bar_label(self.bars,
                  "\{v:.2f\}".format(v=v) for v in vals, padding=1)
plt.draw()

def left(self, event):
    self.loc_filt.go("left")
    self.draw_dist()

def right(self, event):
    self.loc_filt.go("right")
    self.draw_dist()

def door(self, event):
    self.loc_filt.observe({"door":True})
    self.draw_dist()

def no_door(self, event):
    self.loc_filt.observe({"door":False})
    self.draw_dist()

def reset(self, event):
    self.loc_filt.state_dist = {i:1/16 for i in range(16)}
    self.draw_dist()

# sl = Show_Localization(hmm_16pos)
```

9.9.3 Particle Filtering for HMMs

In this implementation a particle is just a state. If you want to do some form of smoothing, a particle should probably be a history of states. This maintains, particles, an array of states, weights an array of (non-negative) real numbers, such that weights[i] is the weight of particles[i].

```python
from display import Displayable
from probStochSim import resample

class HMMpaticleFilter(Displayable):
    def __init__(self, hmm, number_particles=1000):
        self.hmm = hmm
        self.particles = [sample_one(hmm.indist)
                          for i in range(number_particles)]
        self.weights = [1 for i in range(number_particles)]

def filter(self, obsseq):
    """returns the state distribution following the sequence of
    observations in obsseq using particle filtering."
```

http://aipython.org  Version 0.9.4  March 2, 2022
Note that it first advances time. This is what is required if it is called after previous filtering. If that is not what is wanted initially, do an observe first.

```python
for obs in obsseq:
    self.advance()  # advance time
    self.observe(obs)  # observe
    self.resample_particles()
    self.display(2, "After observing", str(obs),
               "state distribution:",
               self.histogram(self.particles))
    self.display(1, "Final state distribution:",
               self.histogram(self.particles))
return self.histogram(self.particles)
```

```python
def advance(self):
    """advance to the next time. This assumes that all of the weights are 1.""
    self.particles = [sample_one(self.hmm.trans[st])
                       for st in self.particles]

def observe(self, obs):
    """reweighs the particles to incorporate observations obs""
    for i in range(len(self.particles)):
        for obv in obs:
            if obs[obv]:
                self.weights[i] *= self.hmm.pobs[obv][self.particles[i]]
            else:
                self.weights[i] *=
                                 1-self.hmm.pobs[obv][self.particles[i]]

def histogram(self, particles):
    """returns list of the probability of each state as represented by
the particles""
    tot=0
    hist = {st: 0.0 for st in self.hmm.states}
    for (st,wt) in zip(particles, self.weights):
        hist[st] += wt
        tot += wt
    return {st:hist[st]/tot for st in hist}

def resample_particles(self):
    """resamples to give a new set of particles.""
    self.particles = resample(self.particles, self.weights,
                                 len(self.particles))
    self.weights = [1] * len(self.particles)
```

The following are some queries for `hmm1`.
Exercise 9.7 A form of importance sampling can be obtained by not resampling. Is it better or worse than particle filtering? Hint: you need to think about how they can be compared. Is the comparison different if there are more states than particles?

Exercise 9.8 Extend the particle filtering code to continuous variables and observations. In particular, suppose the state transition is a linear function with Gaussian noise of the previous state, and the observations are linear functions with Gaussian noise of the state. You may need to research how to sample from a Gaussian distribution.

9.9.4 Generating Examples

The following code is useful for generating examples.

```python
def simulate(hmm, horizon):
    """returns a pair of (state sequence, observation sequence) of length horizon.
    for each time t, the agent is in state_sequence[t] and observes observation_sequence[t]
    ""
    state = sample_one(hmm.indist)
    obsseq=[]
    stateseq=[]
    for time in range(horizon):
        stateseq.append(state)
        newobs =
        {obs:sample_one({0:1-hmm.pobs[obs][state],1:hmm.pobs[obs][state]})
            for obs in hmm.obsvars}
        obsseq.append(newobs)
        state = sample_one(hmm.trans[state])
    return stateseq,obsseq

def simobs(hmm, stateseq):
```
9. Reasoning Under Uncertainty

```python
"""returns observation sequence for the state sequence"
obsseq=[]
for state in stateseq:
    newobs =
        {obs:sample_one({0:1-hmm.pobs[obs][state],1:hmm.pobs[obs][state]})
        for obs in hmm.obsvars}
obsseq.append(newobs)
return obsseq

def create_eg(hmm,n):
    """Create an annotated example for horizon n"
    seq,obs = simulate(hmm,n)
    print("True state sequence:",seq)
    print("Sequence of observations:
",obs)
    hmmfilter = HMMVEfilter(hmm)
    dist = hmmfilter.filter(obs)
    print("Resulting distribution over states:
",dist)
```

9.10 Dynamic Belief Networks

A **dynamic belief network (DBN)** is a belief network that extends in time.

There are a number of ways that reasoning can be carried out in a DBN, including:

- Rolling out the DBN for some time period, and using standard belief network inference. The latest time that needs to be in the rolled out network is the time of the latest observation or the time of a query (whichever is later). This allows us to observe any variables at any time and query any variables at any time. This is covered in Section 9.10.2.

- An unrolled belief network may be very large, and we might only be interested in asking about “now”. In this case we can just representing the variables “now”. In this approach we can observe and query the current variables. We can then move to the next time. This does not allow for arbitrary historical queries (about the past or the future), but can be much simpler. This is covered in Section 9.10.3.

9.10.1 Representing Dynamic Belief Networks

To specify a DBN, think about the distribution *now*. *Now* will be represented as time 1. Each variable will have a corresponding previous variable; these will be created together.

A dynamic belief network consists of:

- A set of features. A variable is a feature-time pair.
9.10. Dynamic Belief Networks

- An initial distribution over the features "now" (time 1). This is a belief network with all variables being time 1 variables.

- A specification of the dynamics. We define the how the variables now (time 1) depend on variables now and the previous time (time 0), in such a way that the graph is acyclic.

```python
from probVariables import Variable
from probGraphicalModels import GraphicalModel, BeliefNetwork
from probFactors import Prob, Factor, CPD
from probVE import VE
from display import Displayable
from utilities import dict_union

class DBNvariable(Variable):
    """A random variable that incorporates the stage (time)"
    A variable can have both a name and an index. The index defaults to 1.
    """
    def __init__(self,name,domain=[False,True],index=1):
        Variable.__init__(self,f"{name}_{index}",domain)
        self.basename = name
        self.domain = domain
        self.index = index
        self.previous = None

    def __lt__(self,other):
        if self.name != other.name:
            return self.name<other.name
        else:
            return self.index<other.index

    def __gt__(self,other):
        return other<self

def variable_pair(name,domain=[False,True]):
    """returns a variable and its predecessor. This is used to define 2-stage DBNs"""
    if name is X, it returns the pair of variables X_prev,X_now"
    var_now = DBNvariable(name,domain,index='now')
    var_prev = DBNvariable(name,domain,index='prev')
    var_now.previous = var_prev
    return var_prev, var_now

A FactorRename is a factor that is the result renaming the variables in the factor. It takes a factor, fac, and a {new : old} dictionary, where new is the name of a variable in the resulting factor and old is the corresponding name in fac. This assumes that the all variables are renamed.

http://aipython.org
```
The following class renames the variables of a conditional probability distribution. It is used for template models (e.g., dynamic decision networks or relational models).
9.10. Dynamic Belief Networks

```python
self.vars_prev = [v.previous for v in vars_now]
self.transition_factors = transition_factors
self.init_factors = init_factors
self.var_index = {}  # var_index[v] is the index of variable v
for i, v in enumerate(vars_now):
    self.var_index[v] = i

Here is a 3 variable DBN:

```probDBN.py — (continued)```

```python
A0, A1 = variable_pair("A", domain=[False, True])
B0, B1 = variable_pair("B", domain=[False, True])
C0, C1 = variable_pair("C", domain=[False, True])

# dynamics
pc = Prob(C1, [B1, C0], [[[0.03, 0.97], [0.38, 0.62]],
                           [[0.23, 0.77], [0.78, 0.22]]])
pb = Prob(B1, [A0, A1], [[[0.5, 0.5], [0.77, 0.23]],
                           [[0.4, 0.6], [0.83, 0.17]]])
pa = Prob(A1, [A0, B0], [[[0.1, 0.9], [0.65, 0.35]],
                           [[0.3, 0.7], [0.8, 0.2]]])

# initial distribution
pa0 = Prob(A1, [], [0.9, 0.1])
pb0 = Prob(B1, [A1], [[0.3, 0.7], [0.8, 0.2]])
pc0 = Prob(C1, [], [0.2, 0.8])

dbn1 = DBN("Simple DBN", [A1, B1, C1], [pa, pb, pc], [pa0, pb0, pc0])
```

Here is the animal example:

```python
from probHMM import closeMic, farMic, midMic, sm, mmc, sc, mcm, mcc

Pos_0, Pos_1 = variable_pair("Position", domain=[0, 1, 2, 3])
Mic1_0, Mic1_1 = variable_pair("Mic1")
Mic2_0, Mic2_1 = variable_pair("Mic2")
Mic3_0, Mic3_1 = variable_pair("Mic3")

# conditional probabilities - see hmm for the values of sm, mmc, etc
ppos = Prob(Pos_1, [Pos_0],
            [[sm, mmc, mmc, mmc], # was in middle
             [mcm, sc, mmc, mcc], # was in corner 1
             [mcm, mcc, sc, mcc], # was in corner 2
             [mcm, mcc, mcc, sc]])  # was in corner 3
pm1 = Prob(Mic1_1, [Pos_1], [[1-midMic, midMic], [1-closeMic, closeMic],
                               [1-farMic, farMic], [1-farMic, farMic]])
pm2 = Prob(Mic2_1, [Pos_1], [[1-midMic, midMic], [1-farMic, FarMic],
                               [1-closeMic, closeMic], [1-farMic, farMic]])
pm3 = Prob(Mic3_1, [Pos_1], [[1-midMic, midMic], [1-farMic, FarMic],
                               [1-farMic, farMic], [1-closeMic, closeMic]])
ipos = Prob(Pos_1, [], [0.25, 0.25, 0.25, 0.25])

dbn_an = DBN("Animal DBN", [Pos_1, Mic1_1, Mic2_1, Mic3_1],
             [ppos, pm1, pm2, pm3],
             [ipos, pm1, pm2, pm3])
```

http://aipython.org  Version 0.9.4  March 2, 2022
9.10.2 Unrolling DBNs

```python
class BNfromDBN(BeliefNetwork):
    """Belief Network unrolled from a dynamic belief network
    ""
    def __init__(self, dbn, horizon):
        """dbn is the dynamic belief network being unrolled
        horizon>0 is the number of steps (so there will be horizon+1
        variables for each DBN variable.
        ""
        self.name2var = {var.basename:
            [DBNvariable(var.basename, var.domain, index) for index in
             range(horizon+1)]
            for var in dbn.vars_now}
        self.display(1,f"name2var={self.name2var}")
        variables = {v for vs in self.name2var.values() for v in vs}
        self.display(1,f"variables={variables}")
        bnfactors = {CPDrename(fac,{self.name2var[var.basename][0]:var
            for var in fac.variables})
            for fac in dbn.init_factors}
        bnfactors |=
            {CPDrename(fac,dict_union({self.name2var[var.basename][i]:var
                for var in fac.variables if
                var.index=='prev'})
            , {self.name2var[var.basename][i+1]:var
                for var in fac.variables if
                var.index=='now'})
            for fac in dbn.transition_factors
            for i in range(horizon)}
        self.display(1,f"bnfactors={bnfactors}")
        BeliefNetwork.__init__(self, dbn.title, variables, bnfactors)
```

Here are two examples. Note that we need to use `bn.name2var['B'][2]` to get the variable B2 (B at time 2).

```python
# Try
# from probRC import ProbRC
# bn = BNfromDBN(dbn1,2) # construct belief network
# drc = ProbRC(bn) # initialize recursive conditioning
# B2 = bn.name2var['B'][2]
# drc.query(B2) #P(B2)
# drc.query(bn.name2var['B'][1],{bn.name2var['B'][0]:True,bn.name2var['C'][1]:False})
# P(B1|B0,C1)
```

9.10.3 DBN Filtering

If we only wanted to ask questions about the current state, we can save space by forgetting the history variables.

[http://aipython.org](http://aipython.org)
class DBNVEFilter(VE):
    def __init__(self, dbn):
        self.dbn = dbn
        self.current_factors = dbn.init_factors
        self.current_obs = {}

    def observe(self, obs):
        """updates the current observations with obs.
        obs is a variable:value dictionary where variable is a current
        variable.  """
        assert all(self.current_obs[var] == obs[var] for var in obs
        if var in self.current_obs), "inconsistent current
        observations"
        self.current_obs.update(obs)  # note 'update' is a dict method

    def query(self, var):
        """returns the posterior probability of current variable var""
        return VE(GraphicalModel(self.dbn.title, self.dbn.vars_now, self.current_factors)).query(var, self.current_obs)

    def advance(self):
        """advance to the next time""
        prev_factors = [self.make_previous(fac) for fac in self.current_factors]
        prev_obs = {var.previous: val for var, val in self.current_obs.items()}
        two_stage_factors = prev_factors + self.dbn.transition_factors
        self.current_factors = self.elim_vars(two_stage_factors, self.dbn.vars_prev, prev_obs)
        self.current_obs = {}

    def make_previous(self, fac):
        """Creates new factor from fac where the current variables in fac
        are renamed to previous variables.  """
        return FactorRename(fac, {var.previous: var for var in fac.variables})

    def elim_vars(self, factors, vars, obs):
        for var in vars:
            if var in obs:
                factors = [self.project_observations(fac, obs) for fac in factors]
            else:
                factors = self.eliminate_var(factors, var)
        return factors

Example queries:

http://aipython.org
9. Reasoning Under Uncertainty

# probDBN.py — (continued)

```python
# df = DBNVEfilter(dbn1)
# df.observe({B1:True}); df.advance(); df.observe({C1:False})
# df.query(B1) #P(B1|B0,C1)
# df.advance(); df.query(B1)
# dfa = DBNVEfilter(dbn_an)
# dfa.observe({Mic1_1:0, Mic2_1:1, Mic3_1:1})
# dfa.advance()
# dfa.observe({Mic1_1:1, Mic2_1:0, Mic3_1:1})
# dfa.query(Pos_1)
```

9.11 Causal Models

A causal model can answer “do” questions.

The following adds the queryDo method to the InferenceMethod class, so it can be used with any inference method.

```python
from probGraphicalModels import InferenceMethod, BeliefNetwork
from probFactors import CPD, ConstantCPD

def queryDo(self, qvar, obs={}, do={}):
    assert isinstance(self.gm, BeliefNetwork), "Do only applies to belief networks"
    if do=={}:
        return self.query(qvar, obs)
    else:
        newfacs = ({f for (ch,f) in self.gm.var2cpt.items() if ch not in do} |
                    {ConstantCPD(v,c) for (v,c) in do.items()})
        self.modBN = BeliefNetwork(self.gm.title+"(mod)",
                                    self.gm.variables, newfacs)
        oldBN, self.gm = self.gm, self.modBN
        result = self.query(qvar, obs)
        self.gm = oldBN # restore original
        return result

InferenceMethod.queryDo = queryDo
```

```python
from probRC import ProbRC

from probGraphicalModels import bn_sprinkler, Season, Sprinkler, Rained,
                              Grass_wet, Grass_shiny, Shoes_wet, bn_sprinkler_soff
bn_sprinklerv = ProbRC(bn_sprinkler)
# bn_sprinklerv.queryDo(Shoes_wet)
# bn_sprinklerv.queryDo(Shoes_wet,obs={Sprinkler:"off"})
# bn_sprinklerv.queryDo(Shoes_wet,do={Sprinkler:"off"})
```

http://aipython.org  Version 0.9.4  March 2, 2022
9.11. Causal Models

```python
from probVariables import Variable
from probFactors import Prob
from probGraphicalModels import boolean

Drug_Prone = Variable("Drug_Prone", boolean, position=(0.1,0.5))
Takes_Marijuana = Variable("Takes_Marijuana", boolean, position=(0.1,0.5))
Side_Effects = Variable("Side_Effects", boolean, position=(0.1,0.5))
Takes_Hard_Drugs = Variable("Takes_Hard_Drugs", boolean,
    position=(0.9,0.5))

p_dp = Prob(Drug_Prone, [], [0.8, 0.2])
p_tm = Prob(Takes_Marijuana, [Drug_Prone], [[0.98, 0.02], [0.2, 0.8]])
p_be = Prob(Side_Effects, [Takes_Marijuana], [[1, 0], [0.4, 0.6]])
p_thd = Prob(Takes_Hard_Drugs, [Side_Effects, Drug_Prone],
    # Drug_Prone=False Drug_Prone=True
    [[[0.999, 0.001], [0.6, 0.4]], # Side_Effects=False
    [[0.9999, 0.00001], [0.995, 0.005]]]) # Side_Effects=True
drugs = BeliefNetwork("Gateway Drugs",
    [Drug_Prone,Takes_Marijuana,Side_Effects,Takes_Hard_Drugs],
    [p_dp, p_tm, p_be, p_thd])
drugsq = ProbRC(drugs)
# drugsq.queryDo(Takes_Hard_Drugs)
# drugsq.queryDo(Takes_Hard_Drugs, obs = {Takes_Marijuana: True})
# drugsq.queryDo(Takes_Hard_Drugs, obs = {Takes_Marijuana: False})
# drugsq.queryDo(Takes_Hard_Drugs, do = {Takes_Marijuana: True})
# drugsq.queryDo(Takes_Hard_Drugs, do = {Takes_Marijuana: False})
```
10.1 Decision Networks

The decision network code builds on the representation for belief networks of Chapter 9.

We first allow for factors that define the utility. Here the utility is a function of the variables in \( \texttt{vars} \). In a utility table the utility is defined in terms of \( \texttt{a} \), a list that enumerates the values as in Section 9.3.3.

A decision variable is a like a random variable with a string name, and a domain, which is a list of possible values. The decision variable also includes the parents, a list of the variables whose value will be known when the decision is made. It also includes a potion, which is only used for plotting.
A decision network is a graphical model where the variables can be random variables or decision variables. Among the factors we assume there is one utility factor.

The split order ensures that the parents of a decision node are split before the decision node, and no other variables (if that is possible).
10.1. Decision Networks

Umbrella Decision Network

for par in self.utility_factor.variables:
    ax.annotate("Utility", par.position,
                xytext=self.utility_factor.position,
                arrowprops={'arrowstyle':'<-'},bbox=dict(boxstyle="sawtooth,pad=1.0",ha='center')

for var in reversed(self.topological_sort()):
    if isinstance(var,DecisionVariable):
        bbox = dict(boxstyle="square,pad=1.0",color="green")
    else:
        bbox = dict(boxstyle="round4,pad=1.0,rounding_size=0.5")
    if self.var2parents[var]:
        for par in self.var2parents[var]:
            ax.annotate(var.name, par.position, xytext=var.position,
                        arrowprops={'arrowstyle':'<-'},bbox=bbox,ha='center')
    else:
        x,y = var.position
        plt.text(x,y,var.name,bbox=bbox,ha='center')

10.1.1 Example Decision Networks

Umbrella Decision Network

Here is a simple “umbrella” decision network. The output of umbrella_dn.show() is shown in Figure 10.1.

```python
Weather = Variable("Weather", ["NoRain", "Rain"], position=(0.5,0.8))
Forecast = Variable("Forecast", ["Sunny", "Cloudy", "Rainy"],
                       position=(0,0.4))
# Each variant uses one of the following:
```

http://aipython.org Version 0.9.4 March 2, 2022
10. Planning with Uncertainty

Umbrella = DecisionVariable("Umbrella", ["Take", "Leave"], {Forecast},
    position=(0.5,0))

p_weather = Prob(Weather, [], [0.7, 0.3])

p_forecast = Prob(Forecast, [Weather], [[0.7, 0.2, 0.1], [0.15, 0.25,
    0.6]])

umb_utility = UtilityTable([Weather, Umbrella], [[20, 100], [70, 0]],
    position=(1,0.4))

umbrella_dn = DecisionNetwork("Umbrella Decision Network",
    {Weather, Forecast, Umbrella},
    {p_weather, p_forecast, umb_utility})

The following is a variant with the umbrella decision having 2 parents; nothing
else has changed. This is interesting because one of the parents is not needed;
if the agent knows the weather, it can ignore the forecast.

umbrella2p = DecisionVariable("Umbrella", ["Take", "Leave"], {Forecast,
    Weather}, position=(0.5,0))

umb_utility2p = UtilityTable([Weather, Umbrella2p], [[20, 100], [70, 0]],
    position=(1,0.4))

umbrella_dn2p = DecisionNetwork("Umbrella Decision Network (extra arc)",
    {Weather, Forecast, Umbrella2p},
    {p_weather, p_forecast, umb_utility2p})

Fire Decision Network

The fire decision network of Figure 10.2 (showing the result of fire_dn.show())
is represented as:

boolean = [False, True]

Alarm = Variable("Alarm", boolean, position=(0.25,0.633))

Fire = Variable("Fire", boolean, position=(0.5,0.9))

Leaving = Variable("Leaving", boolean, position=(0.25,0.366))

Report = Variable("Report", boolean, position=(0.25,0.1))

Smoke = Variable("Smoke", boolean, position=(0.75,0.633))

Tamper = Variable("Tamper", boolean, position=(0,0.9))

See_Sm = Variable("See_Sm", boolean, position=(0.75,0.366) )

Chk_Sm = DecisionVariable("Chk_Sm", boolean, {Report}, position=(0.5,
    0.366))

Call = DecisionVariable("Call", boolean,{See_Sm,Chk_Sm,Report},
    position=(0.75,0.1))

f_ta = Prob(Tamper, [], [0.98,0.02])

f_fi = Prob(Fire, [], [0.99,0.01])

f_sm = Prob(Smoke,[Fire],[[0.99,0.01],[0.1,0.9]])

f_al = Prob(Alarm,[Fire,Tamper],[[0.9999, 0.0001], [0.15, 0.85]], [[0.01,
    0.99], [0.5, 0.5]])
### 10.1. Decision Networks

#### Fire Decision Network

![Fire Decision Network Diagram](diagram.png)

- **Tamper** ➝ **Fire** ➝ **Alarm** ➝ **Leaving** ➝ **Report** ➝ **Call**

#### Variables and Utilities

- \( f_{lv} = \text{Prob}(\text{Leaving}, [\text{Alarm}], [[0.999, 0.001], [0.12, 0.88]]) \)
- \( f_{re} = \text{Prob}(\text{Report}, [\text{Leaving}], [[0.99, 0.01], [0.25, 0.75]]) \)
- \( f_{ss} = \text{Prob}(\text{See_Sm}, [\text{Chk_Sm, Smoke}], [[[1,0],[1,0]],[[1,0],[0,1]]]) \)
- \( \text{ut} = \text{UtilityTable}([\text{Chk_Sm, Fire, Call}], [[[0,-200],[-5000,-200]],[[-20,-220],[-5020,-220]]], \text{position}=(1,0.633)) \)

**Fire Decision Network**

The following is the representation of the cheating decision of Figure 10.3. Note that we keep the names of the variables short (less than 8 characters) so that the tables look good when printed.

#### Cheating Decision Network

The following is the representation of the cheating decision of Figure 10.3. Note that we keep the names of the variables short (less than 8 characters) so that the tables look good when printed.

- **grades** = ['A', 'B', 'C', 'F']
- **Watched** = Variable("Watched", boolean, position=(0,0.9))
- **Caught1** = Variable("Caught1", boolean, position=(0.2,0.7))
- **Caught2** = Variable("Caught2", boolean, position=(0.6,0.7))
132 | Punish = Variable("Punish", ["None","Suspension","Recorded"],
133      position=(0.8,0.9))
134 | Grade_1 = Variable("Grade_1", grades, position=(0.2,0.3))
135 | Grade_2 = Variable("Grade_2", grades, position=(0.6,0.3))
136 | Fin_Grd = Variable("Fin_Grd", grades, position=(0.8,0.1))
137 | Cheat_1 = DecisionVariable("Cheat_1", boolean, set(), position=(0,0.5))
138 | #no parents
139 | Cheat_2 = DecisionVariable("Cheat_2", boolean, {Cheat_1,Caught1},
140      position=(0,0.5))
141 | p_wa = Prob(Watched,[[0.7, 0.3]])
142 | p_cc1 = Prob(Caught1,[Watched,Cheat_1],[[1.0, 0.0], [0.9, 0.1]], [[1.0,
143      0.0], [0.5, 0.5]])
144 | p_cc2 = Prob(Caught2,[Watched,Cheat_2],[[1.0, 0.0], [0.9, 0.1]], [[1.0,
145      0.0], [0.5, 0.5]])
146 | p_pun = Prob(Punish,[Caught1,Caught2],[[1.0, 0.0, 0.0], [0.5, 0.4, 0.1]],
147      [[0.6, 0.2, 0.2], [0.2, 0.5, 0.3]])
148 | p_gr1 = Prob(Grade_1,[Cheat_1], [('A':0.2, 'B':0.3, 'C':0.3, 'D': 0.2),
149       ('A':0.5, 'B':0.3, 'C':0.2, 'D':0.0)])
150 | p_gr2 = Prob(Grade_2,[Cheat_2], [('A':0.2, 'B':0.3, 'C':0.3, 'D': 0.2),
151       ('A':0.5, 'B':0.3, 'C':0.2, 'D':0.0)])
152 | p_fg = Prob(Fin_Grd,[Grade_1,Grade_2],
153       {'A':('A':1.0, 'B':0.0, 'C': 0.0, 'D':0.0),
154       'B': ('A':0.5, 'B':0.5, 'C': 0.0, 'D':0.0),
155       'C': ('A':0.0, 'B':0.0, 'C': 0.5, 'D':0.0),
156       'D': ('A':0.0, 'B':0.0, 'C': 0.0, 'D':1.0))

Figure 10.3: Cheating Decision Network
10.1. Decision Networks

'S3': {'A': 0.25, 'B': 0.5, 'C': 0.25, 'D': 0.25},
'S2': {'A': 0.5, 'B': 0.5, 'C': 0.0, 'D': 0.0},
'S1': {'A': 0.0, 'B': 1, 'C': 0.0, 'D': 0.0},
'S0': {'A': 0.0, 'B': 0.5, 'C': 0.5, 'D': 0.0},
'D0': {'A': 0.0, 'B': 0.25, 'C': 0.5, 'D': 0.25},
'D1': {'A': 0.0, 'B': 0.0, 'C': 0.1, 'D': 0.0},
'D2': {'A': 0.0, 'B': 0.0, 'C': 0.5, 'D': 0.5},
'
utc = UtilityTable([Punish, Fin_Grd], ['None': {'A': 100, 'B': 90, 'C': 70, 'D': 50},
'Suspension': {'A': 40, 'B': 20, 'C': 10, 'D': 0},
'Recorded': {'A': 70, 'B': 60, 'C': 40, 'D': 20}, position=(1.0, 5))

cheat_dn = DecisionNetwork("Cheat Decision",
{Punish, Caught2, Watched, Fin_Grd, Grade_2, Grade_1, Cheat_2, Caught1, Cheat_1},
{p_wa, p_cc1, p_cc2, p_pun, p_gr1, p_gr2, p_fg, utc})

Chain of 3 decisions

The following example is a finite-stage fully-observable Markov decision process with a single reward (utility) at the end. It is interesting because the parents do not include all predecessors. The methods we use will work without change on this, even though the agent does not condition on all of its previous observations and actions. The output of ch3.show() is shown in Figure 10.4.
10. Planning with Uncertainty

Figure 10.4: A decision network that is a chain of 3 decisions

An instance of a `RC_DN` object takes in a decision network. The query method uses recursive conditioning to compute the expected utility of the optimal policy. `self.opt_policy` becomes the optimal policy.

```
import math
from probGraphicalModels import GraphicalModel, InferenceMethod
from probFactors import Factor
from utilities import dict_union
from probRC import connected_components

class RC_DN(InferenceMethod):
    """The class that queries graphical models using recursive conditioning
```
10.1. Decision Networks

```
10.1. Decision Networks

gm is graphical model to query

""

def __init__(self, gm=None):
    self.gm = gm
    self.cache = {((frozenset(), frozenset())), 1}
    # self.max_display_level = 3

def optimize(self, split_order=None):
    """computes expected utility, and creates optimal decision
    functions, where
    elim_order is a list of the non-observed non-query variables in gm
    """
    if split_order == None:
        split_order = self.gm.split_order()
    self.opt_policy = {}
    return self.rc({}, self.gm.factors, split_order)

    The following us the simplest search-based algorithm. It is exponential in
the number of variables, so is not very useful. However, it is simple, and useful
to understand before looking at the more complicated algorithm. Note that the
above code does not call rc0; you will need to change the self.rc to self.rc0
in above code to use it.

---

def rc0(self, context, factors, split_order):
    """simplest search algorithm"
    self.display(2, "calling rc0", (context, factors), "with
    SO", split_order)
    if not factors:
        return 1
    elif to_eval := {fac for fac in factors if
        fac.can_evaluate(context)}:
        self.display(3, "rc0 evaluating factors", to_eval)
        val = math.prod(fac.get_value(context) for fac in to_eval)
        return val * self.rc0(context, factors-to_eval, split_order)
    else:
        var = split_order[0]
        self.display(3, "rc0 branching on", var)
        if isinstance(var, DecisionVariable):
            assert set(context) <= set(var.parents),
            f"cannot optimize
            {var} in context {context}"
            maxres = -math.inf
            for val in var.domain:
                self.display(3, "In rc0, branching on", var, ",", val)
                newres = self.rc0(dict_union({var: val}, context),
                    factors, split_order[1:])
                if newres > maxres:
                    maxres = newres
                    theval = val

http://aipython.org Version 0.9.4 March 2, 2022
```
234

10. Planning with Uncertainty

```python
self.opt_policy[frozenset(context.items())] = (var,theval)
return maxres
else:
total = 0
for val in var.domain:
    total += self.rc0(dict_union({var:val},context),
                        factors, split_order[1:])
self.display(3, "rc0 branching on", var,"returning", total)
return total
```

We can combine the optimization for decision networks above, with the improvements of recursive conditioning used for graphical models (Section 9.6, page 188).

```python
def rc(self, context, factors, split_order):
    """ returns the number \( \sum_{\text{split_order}} \prod_{\text{factors}} \) given
    assignments in context
    context is a variable:value dictionary
    factors is a set of factors
    split_order is a list of variables in factors that are not in
    context
    """
    self.display(3,"calling rc",(context,factors))
    ce = (frozenset(context.items()), frozenset(factors)) # key for the
    cache entry
    if ce in self.cache:
        self.display(2,"rc cache lookup",(context,factors))
        return self.cache[ce]
    # if not factors: # no factors; needed if you don't have forgetting
    # and caching
    # return 1
    elif vars_not_in_factors := {var for var in context
                                if not any(var in fac.variables
                                for fac in factors)):
        # forget variables not in any factor
        self.display(3,"rc forgetting variables", vars_not_in_factors)
        return self.rc({key:val for (key,val) in context.items()
                        if key not in vars_not_in_factors},
                        factors, split_order)
    elif to_eval := {fac for fac in factors if
        fac.can_evaluate(context)}:
        # evaluate factors when all variables are assigned
        self.display(3,"rc evaluating factors",to_eval)
        val = math.prod(fac.get_value(context) for fac in to_eval)
        if val == 0:
            return 0
    else:
        return val * self.rc(context, {fac for fac in factors if fac
                                        not in to_eval}, split_order)
```

http://aipython.org Version 0.9.4 March 2, 2022
elif len(comp := connected_components(context, factors,
    split_order)) > 1:
    # there are disconnected components
    self.display(2, "splitting into connected components"); comp
    return(math.prod(self.rc(context,f,eo) for (f,eo) in comp))
else:
    assert split_order, f"split_order empty rc(({context}),{factors})"
    var = split_order[0]
    self.display(3, "rc branching on", var)
    if isinstance(var,DecisionVariable):
        assert set(context) <= set(var.parents), f"cannot optimize 
            {var} in context {context}"
        maxres = -math.inf
        for val in var.domain:
            self.display(3, "In rc, branching on", var,"=",val)
            newres = self.rc(dict_union({var:val},context), factors,
                split_order[1:])
            if newres > maxres:
                maxres = newres
                theval = val
        self.opt_policy[frozenset(context.items())] = (var,theval)
        self.cache[ce] = maxres
        return maxres
    else:
        total = 0
        for val in var.domain:
            total += self.rc(dict_union({var:val},context), factors,
                split_order[1:])
        self.display(3, "rc branching on", var,"returning", total)
        self.cache[ce] = total
        return total

Here is how to run the optimize the example decision networks:

```python
# Umbrella decision network
#urc = RC_DN(umberella_dn)
#urc.optimize()
#urc.opt_policy

#rc_fire = RC_DN(fire_dn)
#rc_fire.optimize()
#rc_fire.opt_policy

#rc_cheat = RC_DN(cheat_dn)
#rc_cheat.optimize()
#rc_cheat.opt_policy

#rc_ch3 = RC_DN(ch3)
#rc_ch3.optimize()
#rc_ch3.opt_policy
```

http://aipython.org Version 0.9.4 March 2, 2022
10.1.3 Variable elimination for decision networks

VE_DN is variable elimination for decision networks. The method optimize is used to optimize all the decisions. Note that optimize requires a legal elimination ordering of the random and decision variables, otherwise it will give an exception. (A decision node can only be maximized if the variables that are not its parents have already been eliminated.)

```python
from probVE import VE

class VE_DN(VE):
    """Variable Elimination for Decision Networks""
    def __init__(self,dn=None):
        """dn is a decision network""
        VE.__init__(self,dn)
        self.dn = dn

    def optimize(self,elim_order=None,obs={}):
        if elim_order == None:
            elim_order = reversed(self.gm.split_order())
        policy = []
        proj_factors = [self.project_observations(fact,obs)
                        for fact in self.dn.factors]
        for v in elim_order:
            if isinstance(v,DecisionVariable):
                to_max = [fac for fac in proj_factors
                           if v in fac.variables and set(fac.variables) <=
                           v.all_vars]
                assert len(to_max)==1, "illegal variable order
                "+str(elim_order)+" at "+str(v)
                newFac = FactorMax(v, to_max[0])
                policy.append(newFac.decision_fun)
                proj_factors = [fac for fac in proj_factors if fac is not
                           to_max[0]]+[newFac]
                self.display(2,"maximizing",v,"resulting
                factor",newFac.brief() )
                self.display(3,newFac)
            else:
                proj_factors = self.eliminate_var(proj_factors, v)
                assert len(proj_factors)==1,"Should there be only one element of
                proj_factors?"
                value = proj_factors[0].get_value({})
        return value,policy
```

http://aipython.org  Version 0.9.4  March 2, 2022
def __init__(self, dvar, factor):
    """dvar is a decision variable.
    factor is a factor that contains dvar and only parents of dvar
    """
    self.dvar = dvar
    self.factor = factor
    vars = [v for v in factor.variables if v is not dvar]
    Factor.__init__(self, vars)
    self.values = [None]*self.size
    self.decision_fun = FactorDF(dvar, vars, [None]*self.size)

def get_value(self, assignment):
    """lazy implementation: if saved, return saved value, else compute it"
    index = self.assignment_to_index(assignment)
    if self.values[index]:
        return self.values[index]
    else:
        max_val = float("-inf")  # -infinity
        new_asst = assignment.copy()
        for elt in self.dvar.domain:
            new_asst[self.dvar] = elt
            fac_val = self.factor.get_value(new_asst)
            if fac_val>max_val:
                max_val = fac_val
                best_elt = elt
        self.values[index] = max_val
        self.decision_fun.values[index] = best_elt
        return max_val

A decision function is a stored factor.

class FactorDF(TabFactor):
    """A decision function"
    def __init__(self, dvar, vars, values):
        TabStored.__init__(self, vars, values)
        self.dvar = dvar
        self.name = str(dvar)  # Used in printing

Here are some example queries:

# Example queries:
# v,p = VE_DN(fire_dn).optimize(); print(v)
# for df in p: print(df,"\n")
# VE_DN.max_display_level = 3 # if you want to show lots of detail
# v,p = VE_DN(cheat_dn).optimize(); print(v)
# for df in p: print(df,"\n") # print decision functions
10.2 Markov Decision Processes

We will represent a Markov decision process (MDP) directly, rather than using the recursive conditioning or variable elimination code, as we did for decision networks.

```python
from utilities import argmaxd
import random
import matplotlib.pyplot as plt
from matplotlib.widgets import Button, CheckButtons

class MDP(object):
    """A Markov Decision Process. Must define:
    self.states the set (or list) of states
    self.actions the set (or list) of actions
    self.discount a real-valued discount
    """

    def __init__(self, states, actions, discount, init=0):
        self.states = states
        self.actions = actions
        self.discount = discount
        self.initv = self.v = {s:init for s in self.states}
        self.initq = self.q = {s: {a: init for a in self.actions} for s in self.states}

    def P(self, s, a):
        """Transition probability function
        returns a dictionary of {s1:p1} such that P(s1 | s,a)=p1. Other
        probabilities are zero.
        """
        raise NotImplementedError("P") # abstract method

    def R(self, s, a):
        """Reward function R(s,a)
        returns the expected reward for doing a in state s.
        """
        raise NotImplementedError("R") # abstract method
```

Two state partying example (Example 9.27 in Poole and Mackworth [2017]):

```python
from mdpProblem import MDP, GridMDP
class party(MDP):
    """Simple 2-state, 2-Action Partying MDP Example"
    def __init__(self, discount=0.9):
        states = {'healthy', 'sick'}
        actions = {'relax', 'party'}
        MDP.__init__(self, states, actions, discount)
```

http://aipython.org  Version 0.9.4  March 2, 2022
10.2. Markov Decision Processes

```python
def R(self, s, a):
    "R(s,a)"
    return {'healthy': {'relax': 7, 'party': 10},
            'sick': {'relax': 0, 'party': 2}}[s][a]

def P(self, s, a):
    "returns a dictionary of {s1:p1} such that P(s1 | s,a)=p1. Other
    probabilities are zero."
    phealthy = { # P('healthy' | s, a)
        'healthy': {'relax': 0.95, 'party': 0.7},
        'sick': {'relax': 0.5, 'party': 0.1}}[s][a]
    return {'healthy':phealthy, 'sick':1-phealthy}
```

The next example is the tiny game from Example 12.1 and Figure 12.1 of Poole and Mackworth [2017]. The state is represented as \((x,y)\) where \(x\) counts from zero from the left, and \(y\) counts from zero upwards, so the state \((0,0)\) is on the bottom-left state. The actions are \(upC\) for up-careful, and \(upR\) for up-risky.

(Note that \(GridMDP\) is just a type of MDP for which we have methods to show; you can assume it is just MDP here).

```python
class MDPtiny(GridMDP):
    def __init__(self, discount=0.9):
        actions = ['right', 'upC', 'left', 'upR']
        self.x_dim = 2 # x-dimension
        self.y_dim = 3
        states = [(x,y) for x in range(self.x_dim) for y in range(self.y_dim)]
        # for GridMDP
        self.xoff = {'right':0.25, 'upC':0, 'left':-0.25, 'upR':0}
        self.yoff = {'right':0, 'upC':-0.25, 'left':0, 'upR':0.25}
        GridMDP.__init__(self, states, actions, discount)

    def P(self, s, a):
        """return a dictionary of {s1:p1} if P(s1 | s,a)=p1. Other
        probabilities are zero."
        (x,y) = s
        if a == 'right':
            return {(1,y):1}
        elif a == 'upC':
            return {(x,min(y+1,2)):1}
        elif a == 'left':
            if (x,y) == (0,2): return {(0,0):1}
            else: return {(0,y): 1}
        elif a == 'upR':
            if x==0:
                if y<2: return {(x,y):0.1, (x+1,y):0.1, (x,y+1):0.8}
                else: # at (0,2)
                    return {(0,0):0.1, (1,2): 0.1, (0,2): 0.8}
```
elif y < 2:  # x==1
    return {(0,y):0.1, (1,y):0.1, (1,y+1):0.8}
else:  # at (1,2)
    return {(0,2):0.1, (1,2): 0.9}

def R(self,s,a):
    (x,y) = s
    if a == 'right':
        return [0,-1][x]
    elif a == 'upC':
        return [-1,-1,-2][y]
    elif a == 'left':
        if x==0:
            return [-1, -100, 10][y]
        else:
            return 0
    elif a == 'upR':
        return [[-0.1, -10, 0.2],[-0.1, -0.1, -0.9]][x][y]
    # at (0,2) reward is 0.1*10+0.8*(-1)=0.2

Here is the domain of Example 9.28 of Poole and Mackworth [2017]. Here
the state is represented as (x,y) where x counts from zero from the left, and y
counts from zero upwards, so the state (0,0) is on the bottom-left state.
10.2. Markov Decision Processes

```
return (x-1 if x > 0 else x,y)

def P(self,s,a):
    """return a dictionary of {s1:p1} if P(s1 | s,a)=p1. Other probabilities are zero.
    Corners are tricky because different actions result in same state."
    if s in self.fling_states:
        return { (0,0): 0.25, (self.x_dim-1,0):0.25,
            (0,self.y_dim-1):0.25, (self.x_dim-1,self.y_dim-1):0.25 }
    res = dict()
    for ai in self.actions:
        s1 = self.intended_next(s,ai)
        ps1 = 0.7 if ai==a else 0.1
        if s1 in res: # occurs in corners
            res[s1] += ps1
        else:
            res[s1] = ps1
    return res

def R(self,s,a):
    if s in self.rewarding_states:
        return self.rewarding_states[s]
    else:
        (x,y) = s
        rew = 0
        # rewards from crashing:
        if y==0:  # on bottom.
            rew += -0.7 if a == 'down' else -0.1
        if y==self.y_dim-1:  # on top.
            rew += -0.7 if a == 'up' else -0.1
        if x==0:  # on left
            rew += -0.7 if a == 'left' else -0.1
        if x==self.x_dim-1:  # on right.
            rew += -0.7 if a == 'right' else -0.1
    return rew
```

10.2.1 Value Iteration

This implements value iteration.

This uses indexes of the states and actions (not the names). The value function is represented so $v[s]$ is the value of state with index $s$. A $Q$ function is represented so $q[s][a]$ is the value for doing action with index $a$ state with index $s$. Similarly a policy $\pi$ is represented as a list where $\pi[s]$, where $s$ is the index of a state, returns the index of the action.
10. Planning with Uncertainty

"""carries out n iterations of value iteration, updating value
   function self.v
Returns a Q-function, value function, policy
"""
print("calling vi")
assert n>0,"You must carry out at least one iteration of vi. 
n="+str(n)
#v = v0 if v0 is not None else {s:0 for s in self.states}
for i in range(n):
    self.q = {s: {a: self.R(s,a)+self.discount*sum(p1*self.v[s1]
                   for (s1,p1) in
                   self.P(s,a).items())
               for a in self.actions}
           for s in self.states}
    self.v = {s: max(self.q[s][a] for a in self.actions)
               for s in self.states}
    self.pi = {s: argmaxd(self.q[s])
               for s in self.states}
return self.q, self.v, self.pi

The following shows how this can be used.

# Testing value iteration
# Try the following:
# pt = party(discount=0.9)
# pt.vi(1)
# pt.vi(100)
# party(discount=0.99).vi(100)
# party(discount=0.4).vi(100)
# gr = grid()
# gr.show()
# q,v,pi = gr.vi(100)
# q[(7,2)]

10.2.2 Showing Grid MDPs

A GridMDP is a type of MDP where the states are (x,y) positions. It is a
special sort of MDP only because we have methods to show it.

class GridMDP(MDP):
    def __init__(self, states, actions, discount):
        MDP.__init__(self, states, actions, discount)
    def show(self):
        #plt.ion() # interactive
        fig,(self.ax) = plt.subplots()
        plt.subplots_adjust(bottom=0.2)
stepB = Button(plt.axes([0.8,0.05,0.1,0.075]), "step")
stepB.on_clicked(self.on_step)
resetB = Button(plt.axes([0.6,0.05,0.1,0.075]), "reset")
resetB.on_clicked(self.on_reset)
sel.qcheck = CheckButtons(plt.axes([0.2,0.05,0.35,0.075]),
["show q-values","show policy"],
sel.qcheck.on_clicked(self.show_vals)
self.show_vals(None)
plt.show()

def show_vals(self,event):
    self.ax.cla()
    array = [[self.v[(x,y)] for x in range(self.x_dim)]
        for y in range(self.y_dim)]
    self.ax.pcolormesh([x-0.5 for x in range(self.x_dim+1)],
        [x-0.5 for x in range(self.y_dim+1)],
        array, edgecolors='black',cmap='summer')
    # for cmap see
    if self.qcheck.get_status()[1]: # "show policy"
        for (x,y) in self.q:
            maxv = max(self.q[(x,y)][a] for a in self.actions)
            for a in self.actions:
                if self.q[(x,y)][a] == maxv:
                    # draw arrow in appropriate direction
                    self.ax.arrow(x,y,self.xoff[a]*2,self.yoff[a]*2,
                        color='red',width=0.05, head_width=0.2,
                        length_includes_head=True)
    else:
        self.show_v(event)
        self.ax.set_xticks(range(self.x_dim))
        self.ax.set_xticklabels(range(self.x_dim))
        self.ax.set_yticks(range(self.y_dim))
        self.ax.set_yticklabels(range(self.y_dim))
        plt.draw()

def on_step(self,event):
    self.vi(1)
    self.show_vals(event)

def show_v(self,event):
    """show values""
    for (x,y) in self.v:
        self.ax.text(x,y,"{val:.2f}".format(val=self.v[(x,y)]),ha='center')

def show_q(self,event):
    """show q-values""
    for (x,y) in self.q:
10. Planning with Uncertainty

Figure 10.5: Interface for tiny example, after a number of steps. Each rectangle represents a state. In each rectangle are the 4 Q-values for the state. The leftmost number is the for the left action; the rightmost number is for the right action; the upper most is for the upR (up-risky) action and the lowest number is for the upC action. The arrow points to the action(s) with the maximum Q-value.

```
for a in self.actions:
    self.ax.text(x+self.xoff[a],y+self.yoff[a],
        "(val:.2f)".format(val=self.q[(x,y)][a]),ha='center')

def on_reset(self,event):
    self.v = self.initv
    self.q = self.initq
    self.show_vals(event)
```

Figure 10.5 shows the user interface, which can be obtained using `tiny().show()`, resizing it, checking “show q-values” and “show policy”, and clicking “step” a few times.

Figure 10.6 shows the user interface, which can be obtained using `grid().show()`, checking “show q-values” and “show policy”, and clicking “step” a few times.
resizing it, checking “show q-values” and “show policy”, and clicking “step” a few times.

**Exercise 10.1** Computing $q$ before $v$ may seem like a waste of space because we don’t need to store $q$ in order to compute value function or the policy. Change the algorithm so that it loops through the states and actions once per iteration, and only stores the value function and the policy. Note that to get the same results as before, you would need to make sure that you use the previous value of $v$ in the computation not the current value of $v$. Does using the current value of $v$ hurt the algorithm or make it better (in approaching the actual value function)?

### 10.2.3 Asynchronous Value Iteration

This implements asynchronous value iteration, storing $Q$.

A $Q$ function is represented so $q[s][a]$ is the value for doing action with index $a$ state with index $s$.

```python
mdpProblem.py — (continued)

```{definition}
def avi(self, n):
    states = list(self.states)
    actions = list(self.actions)
    for i in range(n):
        s = random.choice(states)
        a = random.choice(actions)
        self.q[s][a] = (self.R(s,a) + self.discount *
                        sum(p1 * max(self.q[s1][a] for a1 in self.actions)
                            for (s1,p1) in self.P(s,a).items()))
    return Q
```

The following shows how `avi` can be used.

```python
# Testing asynchronous value iteration
# Try the following:
# pt = party(discount=0.9)
# pt.avi(10)
# pt.vi(1000)
# gr = grid()
# q = gr.avi(100000)
# q[(7,2)]
```

**Exercise 10.2** Implement value iteration that stores the $V$-values rather than the $Q$-values. Does it work better than storing $Q$? (What might better mean?)

**Exercise 10.3** In asynchronous value iteration, try a number of different ways to choose the states and actions to update (e.g., sweeping through the state-action pairs, choosing them at random). Note that the best way may be to determine...
upper most is for the up action and the lowest number is for the down action. The number is the for the left action; the rightmost number is for the right action; the represents a state. In each rectangle are the 4 Q-values for the state. The leftmost

Figure 10.6: Interface for grid example, after a number of steps. Each rectangle represents a state. In each rectangle are the 4 Q-values for the state. The leftmost number is the for the left action; the rightmost number is for the right action; the upper most is for the up action and the lowest number is for the down action. The arrow points to the action(s) with the maximum Q-value.
which states have had their Q-values change the most, and then update the previous ones, but that is not so straightforward to implement, because you need to find those previous states.
11.1 K-means

The k-means learner maintains two lists that suffice as sufficient statistics to classify examples, and to learn the classification:

- `class_counts` is a list such that `class_counts[c]` is the number of examples in the training set with `class = c`.

- `feature_sum` is a list such that `feature_sum[i][c]` is sum of the values for the `i`th feature `i` for members of class `c`. The average value of the `i`th feature in class `i` is

\[
\frac{\text{feature_sum}[i][c]}{\text{class_counts}[c]}
\]

The class is initialized by randomly assigning examples to classes, and updating the statistics for `class_counts` and `feature_sum`.

```python
class K_means_learner(Learner):
    def __init__(self, dataset, num_classes):
        self.dataset = dataset
        self.num_classes = num_classes
        self.random_initialize()
    
    def random_initialize(self):
```
11. Learning with Uncertainty

```python
# class_counts[c] is the number of examples with class=c
self.class_counts = [0]*self.num_classes
# feature_sum[i][c] is the sum of the values of feature i for class c
self.feature_sum = [[0]*self.num_classes
    for feat in self.dataset.input_features]
for eg in self.dataset.train:
    cl = random.randrange(self.num_classes) # assign eg to random class
    self.class_counts[cl] += 1
    for (ind,feat) in enumerate(self.dataset.input_features):
        self.feature_sum[ind][cl] += feat(eg)
self.num_iterations = 0
self.display(1,"Initial class counts: ",self.class_counts)
```

The distance from (the mean of) a class to an example is the sum, over all features, of the sum-of-squares differences of the class mean and the example value.

```python
def distance(self,cl,eg):
    """distance of the eg from the mean of the class""
    return sum((self.class_prediction(ind,cl)-feat(eg))**2
        for (ind,feat) in enumerate(self.dataset.input_features))
```

```python
def class_prediction(self,feat_ind,cl):
    """prediction of the class cl on the feature with index feat_ind""
    if self.class_counts[cl] == 0:
        return 0 # there are no examples so we can choose any value
    else:
        return self.feature_sum[feat_ind][cl]/self.class_counts[cl]
```

```python
def class_of_eg(self,eg):
    """class to which eg is assigned""
    return (min((self.distance(cl,eg),cl)
        for cl in range(self.num_classes)))[1]
```

One step of k-means updates the class_counts and feature_sum. It uses the old values to determine the classes, and so the new values for class_counts and feature_sum. At the end it determines whether the values of these have changes, and then replaces the old ones with the new ones. It returns an indicator of whether the values are stable (have not changed).

```python
def k_means_step(self):
    """Updates the model with one step of k-means.
    Returns whether the assignment is stable."
```

http://aipython.org
new_class_counts = [0]*self.num_classes  
# feature_sum[i][c] is the sum of the values of feature i for class c  
new_feature_sum = [[0]*self.num_classes  
                 for feat in self.dataset.input_features]
for eg in self.dataset.train:  
    cl = self.class_of_eg(eg)  
    new_class_counts[cl] += 1  
    for (ind,feat) in enumerate(self.dataset.input_features):  
        new_feature_sum[ind][cl] += feat(eg)  
stable = (new_class_counts == self.class_counts) and  
         (self.feature_sum == new_feature_sum)  
self.class_counts = new_class_counts  
self.feature_sum = new_feature_sum  
self.num_iterations += 1  
return stable

def learn(self,n=100):  
    """do n steps of k-means, or until convergence"""  
    i=0  
    stable = False  
    while i<n and not stable:  
        stable = self.k_means_step()  
        i += 1  
        self.display(1,"Iteration",self.num_iterations,  
                     "class counts: ",self.class_counts,"  
                      Stable=",stable)  
    return stable

def show_classes(self):  
    """sorts the data by the class and prints in order.  
    For visualizing small data sets  
    """  
    class_examples = [[] for i in range(self.num_classes)]  
    for eg in self.dataset.train:  
        class_examples[self.class_of_eg(eg)].append(eg)  
    print("Class","Example",sep='\t')  
    for cl in range(self.num_classes):  
        for eg in class_examples[cl]:  
            print(cl,*eg,sep='\t')

def plot_error(self, maxstep=20):  
    """Plots the sum-of-squares error as a function of the number of steps"""  
    plt.ion()  
    plt.xlabel("step")  
    plt.ylabel("Ave sum-of-squares error")  
    train_errors = []  
    if self.dataset.test:
**Exercise 11.1** Change `boolean_features = True` flag to allow for numerical features. K-means assumes the features are numerical, so we want to make non-numerical features into numerical features (using characteristic functions) but we probably don’t want to change numerical features into Boolean.

**Exercise 11.2** If there are many classes, some of the classes can become empty (e.g., try 100 classes with `carbool.csv`). Implement a way to put some examples into a class, if possible. Two ideas are:
11.2 EM

(a) Initialize the classes with actual examples, so that the classes will not start empty. (Do the classes become empty?)
(b) In class prediction, we test whether the code is empty, and make a prediction of 0 for an empty class. It is possible to make a different prediction to “steal” an example (but you should make sure that a class has a consistent value for each feature in a loop).

Make your own suggestions, and compare it with the original, and whichever of these you think may work better.

11.2 EM

In the following definition, a class, \( c \), is a integer in range \([0, num\_classes]\). \( i \) is an index of a feature, so \( \text{feat}[i] \) is the \( i \)th feature, and a feature is a function from tuples to values. \( \text{val} \) is a value of a feature.

A model consists of 2 lists, which form the sufficient statistics:

- \( \text{class\_counts} \) is a list such that \( \text{class\_counts}[c] \) is the number of tuples with \( \text{class} = c \), where each tuple is weighted by its probability, i.e.,

\[
\text{class\_counts}[c] = \sum_{t: \text{class}(t) = c} P(t)
\]

- \( \text{feature\_counts} \) is a list such that \( \text{feature\_counts}[i][\text{val}][c] \) is the weighted count of the number of tuples \( t \) with \( \text{feat}[i](t) = \text{val} \) and \( \text{class}(t) = c \), each tuple is weighted by its probability, i.e.,

\[
\text{feature\_counts}[i][\text{val}][c] = \sum_{t: \text{feat}[i](t) = \text{val} \text{ and } \text{class}(t) = c} P(t)
\]

```python
from learnProblem import Data_set, Learner, Data_from_file
import random
import math
import matplotlib.pyplot as plt

class EM_learner(Learner):
    def __init__(self, dataset, num_classes):
        self.dataset = dataset
        self.num_classes = num_classes
        self.class_counts = None
        self.feature_counts = None
```

The function \textit{em\_step} goes though the training examples, and updates these counts. The first time it is run, when there is no model, it uses random distributions.
11. Learning with Uncertainty

```python
def em_step(self, orig_class_counts, orig_feature_counts):
    """updates the model."""
    class_counts = [0]*self.num_classes
    feature_counts = [{val:[0]*self.num_classes
                      for val in feat.frange}
                     for feat in self.dataset.input_features]
    for tple in self.dataset.train:
        if orig_class_counts: # a model exists
            tpl_class_dist = self.prob(tple, orig_class_counts,
                                       orig_feature_counts)
        else: # initially, with no model, return a random
distribution
            tpl_class_dist = random_dist(self.num_classes)
        for cl in range(self.num_classes):
            class_counts[cl] += tpl_class_dist[cl]
            feature_counts[cl] += tpl_class_dist[cl]
    return class_counts, feature_counts
```

`prob` computes the probability of a class \( c \) for a tuple \( tple \), given the current statistics.

\[
P(c \mid tple) \propto P(c) \prod_i P(X_i=tple(i) \mid c) = \frac{\text{class_counts}[c]}{\text{len}(\text{self.dataset})} \prod_i \frac{\text{feature_counts}[i][\text{feat}(tple)][c]}{\text{class_counts}[c]} \propto \prod_i \text{feature_counts}[i][\text{feat}(tple)][c] / \text{class_counts}[c][\text{feats}-1]
\]

The last step is because `len(self.dataset)` is a constant (independent of \( c \)). `class_counts[c]` can be taken out of the product, but needs to be raised to the power of the number of features, and one of them cancels.

```python
def prob(self, tple, class_counts, feature_counts):
    """returns a distribution over the classes for tuple \( tple \) in the
    model defined by the counts"""
    feats = self.dataset.input_features
    unnorm = [prod(feature_counts[i][feat(tple)][c])
              for (i,feat) in enumerate(feats)]
    /(class_counts[c]**(len(feats)-1))
    for c in range(self.num_classes)
    thesum = sum(unnorm)
    return [un/thesum for un in unnorm]
```

`learn` does \( n \) steps of EM:
11.2. EM

```
def learn(self, n):
    """do n steps of EM"""
    for i in range(n):
        self.class_counts, self.feature_counts =
        self.em_step(self.class_counts,
                     self.feature_counts)

The following is for visualizing the classes. It prints the dataset ordered by the
probability of class c.

```
def show_class(self, c):
    """sorts the data by the class and prints in order.
    For visualizing small data sets
    """
    sorted_data =
    sorted((self.prob(tpl, self.class_counts, self.feature_counts)[c],
            ind,  # preserve ordering for equal
                   probabilities
            tpl),
       for (ind, tpl) in enumerate(self.dataset.train))
    for cc, r, tpl in sorted_data:
        print(cc, *tpl, sep=' '

The following are for evaluating the classes.
The probability of a tuple can be evaluated by marginalizing over the classes:

\[
P(tple) = \sum_c P(c) \prod_i P(X_i=tple(i) \mid c) = \sum_c \frac{cc[c]}{\text{len}(self.dataset)} \prod_i \frac{fc[i][\text{feat}_i(tple)][c]}{cc[c]}
\]

where \(cc\) is the class count and \(fc\) is feature count. \(\text{len}(self.dataset)\) can be distributed out of the sum, and \(cc[c]\) can be taken out of the product:

\[
= \frac{1}{\text{len}(self.dataset)} \sum_c \frac{1}{cc[c]^{\text{#feats}-1}} \prod_i fc[i][\text{feat}_i(tple)][c]
\]

Given the probability of each tuple, we can evaluate the logloss, as the negative of
the log probability:

```
def logloss(self, tple):
    """returns the logloss of the prediction on tple, which is
    -log(P(tple)) based on the current class counts and feature counts
    """
    feats = self.dataset.input_features
    res = 0
    cc = self.class_counts
    fc = self.feature_counts
```

http://aipython.org Version 0.9.4 March 2, 2022
for c in range(self.num_classes):
    res += prod(fc[i][feat(tple)][c]
        for (i,feat) in enumerate(feats))/(cc[c]**(len(feats)-1))
if res>0:
    return -math.log2(res/len(self.dataset.train))
else:
    return float("inf") #infinity

def plot_error(self, maxstep=20):
    """Plots the logloss error as a function of the number of steps""
    plt.ion()
    plt.xlabel("step")
    plt.ylabel("Ave Logloss (bits)")
    train_errors = []
    if self.dataset.test:
        test_errors = []
    for i in range(maxstep):
        self.learn(1)
        train_errors.append(sum(self.logloss(tple) for tple in
            self.dataset.train)/len(self.dataset.train))
    if self.dataset.test:
        test_errors.append(sum(self.logloss(tple) for tple in
            self.dataset.test)/len(self.dataset.test))
    plt.plot(range(1,maxstep+1),train_errors,
        label=str(self.num_classes)+" classes. Training set")
    if self.dataset.test:
        plt.plot(range(1,maxstep+1),test_errors,
            label=str(self.num_classes)+" classes. Test set")
    plt.legend()
    plt.draw()

def prod(L):
    """returns the product of the elements of L""
    res = 1
    for e in L:
        res *= e
    return res

def random_dist(k):
    """generate k random numbers that sum to 1""
    res = [random.random() for i in range(k)]
    s = sum(res)
    return [v/s for v in res]
data = Data_from_file('data/emdata2.csv', num_train=10, target_index=2000)
eml = EM_learner(data,2)
num_iter=2
Exercise 11.3 For the EM data, where there are naturally 2 classes, 3 classes does better on the training set after a while than 2 classes, but worse on the test set. Explain why. Hint: look what the 3 classes are. Use “em3.show_class(i)” for each of the classes $i \in \{0, 3\}$.

Exercise 11.4 Write code to plot the logloss as a function of the number of classes (from 1 to say 15) for a fixed number of iterations. (From the experience with the existing code, think about how many iterations is appropriate.)
Multiagent Systems

12.1 Minimax

Here we consider two-player zero-sum games. Here a player only wins when another player loses. This can be modeled as where there is a single utility which one agent (the maximizing agent) is trying minimize and the other agent (the minimizing agent) is trying to minimize.

12.1.1 Creating a two-player game

```
from display import Displayable

class Node(Displayable):
    """A node in a search tree. It has a
    name a string
    isMax is True if it is a maximizing node, otherwise it is minimizing
    node
    children is the list of children
    value is what it evaluates to if it is a leaf.
    """
    def __init__(self, name, isMax, value, children):
        self.name = name
        self.isMax = isMax
        self.value = value
        self.allchildren = children

    def isLeaf(self):
        """"""returns true of this is a leaf node""
        return self.allchildren is None
```
The following is a representation of a magic-sum game, where players take turns picking a number in the range $[1, 9]$, and the first player to have 3 numbers that sum to 15 wins. Note that this is a syntactic variant of tic-tac-toe or naughts and crosses. To see this, consider the numbers on a magic square (Figure 12.1); 3 numbers that add to 15 correspond exactly to the winning positions.
of tic-tac-toe played on the magic square.

Note that we do not remove symmetries. (What are the symmetries? How do the symmetries of tic-tac-toe translate here?)

```python
class Magic_sum(Node):
    def __init__(self, xmove=True, last_move=None, available=[1,2,3,4,5,6,7,8,9], x=[], o=[]):
        """This is a node in the search for the magic-sum game.
        xmove is True if the next move belongs to X.
        last_move is the number selected in the last move
        available is the list of numbers that are available to be chosen
        x is the list of numbers already chosen by x
        o is the list of numbers already chosen by o
        ""
        self.isMax = self.xmove = xmove
        self.last_move = last_move
        self.available = available
        self.x = x
        self.o = o
        self.allchildren = None #computed on demand
        lm = str(last_move)
        self.name = "start" if not last_move else "o="+lm if xmove else "x="+lm

    def children(self):
        if self.allchildren is None:
            if self.xmove:
                self.allchildren = [
                    Magic_sum(xmove = not self.xmove,
                    last_move = sel,
                    available = [e for e in self.available if e is not sel],
                    x = self.x+[sel],
                    o = self.o)
                    for sel in self.available]
            else:
                self.allchildren = [
                    Magic_sum(xmove = not self.xmove,
                    last_move = sel,
                    available = [e for e in self.available if e is not sel],
```

Figure 12.1: Magic Square
x = self.x,
o = self.o+[sel])
for sel in self.available]
return self.allchildren

def isLeaf(self):
    """A leaf has no numbers available or is a win for one of the players. We only need to check for a win for o if it is currently x's turn, and only check for a win for x if it is o's turn (otherwise it would have been a win earlier)."
    return (self.available == [] or
            (sum_to_15(self.last_move,self.o)
             if self.xmove
             else sum_to_15(self.last_move,self.x)))

def evaluate(self):
    if self.xmove and sum_to_15(self.last_move,self.o):
        return -1
    elif not self.xmove and sum_to_15(self.last_move,self.x):
        return 1
    else:
        return 0

def sum_to_15(last,selected):
    """is true if last, together with two other elements of selected sum to 15."
    return any(last+a+b == 15
                for a in selected if a != last
                for b in selected if b != last and b != a)
### 12.1.2 Minimax and $\alpha$-$\beta$ Pruning

This is a naive depth-first **minimax algorithm**:

```python
from aipython import Node

# Define the minimax function
def minimax(node, depth):
    """returns the value of node, and a best path for the agents"""
    if node.isLeaf():
        return node.evaluate(), None
    elif node.isMax:
        max_score = float("-inf")
        max_path = None
        for C in node.children():
            score, path = minimax(C, depth+1)
            if score > max_score:
                max_score = score
                max_path = C.name, path
        return max_score, max_path
    else:
        min_score = float("inf")
        min_path = None
        for C in node.children():
            score, path = minimax(C, depth+1)
            if score < min_score:
                min_score = score
                min_path = C.name, path
        return min_score, min_path
```

---

http://aipython.org  Version 0.9.4  March 2, 2022
The following is a depth-first minimax with \(\alpha-\beta\) pruning. It returns the value for a node as well as a best path for the agents.

```python
def minimax_alpha_beta(node, alpha, beta, depth=0):
    """node is a Node, alpha and beta are cutoffs, depth is the depth
    returns value, path
    where path is a sequence of nodes that results in the value
    ""
    node.display(2, "*depth,"minimax_alpha_beta("",node.name,"",",alpha,"",",beta,"")"
    best=None  # only used if it will be pruned
    if node.isLeaf():
        node.display(2,"*depth,"returning leaf value",node.evaluate())
        return node.evaluate(),None
    elif node.isMax:
        for C in node.children():
            score,path = minimax_alpha_beta(C, alpha, beta, depth+1)
            if score >= beta:  # beta pruning
                node.display(2,"*depth,"pruned due to beta",beta,"C="",C.name)
                return score, None
            if score > alpha:
                alpha = score
                best = C.name, path
                node.display(2,"*depth,"returning max alpha",alpha,"best",best)
                return alpha,best
    else:
        for C in node.children():
            score,path = minimax_alpha_beta(C, alpha, beta, depth+1)
            if score <= alpha:  # alpha pruning
                node.display(2,"*depth,"pruned due to alpha",alpha,"C="",C.name)
                return score, None
            if score < beta:
                beta=score
                best = C.name,path
                node.display(2,"*depth,"returning min beta",beta,"best="",best)
        return beta,best
```

Testing:

```python
from masProblem import fig10_5, Magic_sum, Node

# Node.max_display_level=2 # print detailed trace
# minimax_alpha_beta(fig10_5, -9999, 9999,0)
# minimax_alpha_beta(Magic_sum(), -9999, 9999,0)

#To see how much time alpha-beta pruning can save over minimax, uncomment the following:
http://aipython.org
```
12.1. Minimax

```python
# import timeit
# timeit.Timer("minimax(Magic_sum(),0)",setup="from __main__ import
# minimax, Magic_sum"
# ).timeit(number=1)
# trace=False
# timeit.Timer("minimax_alpha_beta(Magic_sum(), -9999, 9999,0)",
# setup="from __main__ import minimax_alpha_beta, Magic_sum"
# ).timeit(number=1)
```

http://aipython.org
13.1 Representing Agents and Environments

When the learning agent does an action in the environment, it observes a \((state, reward)\) pair from the environment. The \(state\) is the world state; this is the fully observable assumption.

An RL environment implements a \(do(action)\) method that returns a \((state, reward)\) pair.

---

```python
import random
from display import Displayable
from utilities import flip

class RL_env(Displayable):
    def __init__(self, actions, state):
        self.actions = actions # set of actions
        self.state = state # initial state

    def do(self, action):
        """do action
        returns state, reward
        ""
        raise Not Implemented Error("RL_env.do") # abstract method
```

Here is the definition of the simple 2-state, 2-action party/relax decision.

---

```python
class Healthy_env(RL_env):
    def __init__(self):
        RL_env.__init__(self,"party","relax", "healthy")
```

---
def do(self, action):
    """updates the state based on the agent doing action.
    returns state,reward
    """
    if self.state=="healthy":  
        if action=="party":    
            self.state = "healthy" if flip(0.7) else "sick"
            reward = 10
        else: # action=="relax"
            self.state = "healthy" if flip(0.95) else "sick"
            reward = 7
    else: # self.state=="sick"
        if action=="party":    
            self.state = "healthy" if flip(0.1) else "sick"
            reward = 2
        else:
            self.state = "healthy" if flip(0.5) else "sick"
            reward = 0
    return self.state,reward

13.1.1 Simulating an environment from an MDP

Given the definition for an MDP (page 238), Env_from_MDP takes in an MDP and simulates the environment with those dynamics.

Note that the MDP does not contain enough information to simulate a system, because it loses any dependency between the rewards and the resulting state; here we assume the agent always received the average reward for the state and action.

```python
class Env_from_MDP(RL_env):
    def __init__(self, mdp):
        initial_state = mdp.states[0]
        RL_env.__init__(self,mdp.actions, initial_state)
        self.mdp = mdp
        self.action_index = {action:index for (index,action) in enumerate(mdp.actions)}
        self.state_index = {state:index for (index,state) in enumerate(mdp.states)}
    
def do(self, action):
        """updates the state based on the agent doing action.
        returns state,reward
        """
        action_ind = self.action_index[action]
        state_ind = self.state_index[self.state]
        self.state = pick_from_dist(self.mdp.trans[state_ind][action_ind],
                                     self.mdp.states)
        reward = self.mdp.reward[state_ind][action_ind]
```
13.1. Representing Agents and Environments

![Monster game](image)

**Figure 13.1:** Monster game

```python
return self.state, reward

def pick_from_dist(dist,values):
    
    e.g. pick_from_dist([0.3,0.5,0.2],['a','b','c']) should pick 'a' with probability 0.3, etc.
    
    ran = random.random()
    i=0
    while ran>dist[i]:
        ran -= dist[i]
        i += 1
    return values[i]
```

13.1.2 Simple Game

This is for the game depicted in Figure 13.1

```python
import random
from utilities import flip
from rlProblem import RL_env

class Simple_game_env(RL_env):
    xdim = 5
    ydim = 5
    
vwalls = [(0,3), (0,4), (1,4)] # vertical walls right of these locations
    hwalls = [] # not implemented
    crashed_reward = -1
```

[http://aipython.org](http://aipython.org)  Version 0.9.4  March 2, 2022
prize_locs = [(0,0), (0,4), (4,0), (4,4)]
prize_appears_prob = 0.3
prize_reward = 10

monster_locs = [(0,1), (1,1), (2,3), (3,1), (4,2)]
monster_appears_prob = 0.4
monster_reward_when_damaged = -10
repair_stations = [(1,4)]

actions = ["up", "down", "left", "right"]

def __init__(self):
    # State:
    self.x = 2
    self.y = 2
    self.damaged = False
    self.prize = None
    # Statistics
    self.number_steps = 0
    self.total_reward = 0
    self.min_reward = 0
    self.min_step = 0
    self.zero_crossing = 0
    RL_env.__init__(self, Simple_game_env.actions,
                    (self.x, self.y, self.damaged, self.prize))
    self.display(2,"","Step","Tot Rew","Ave Rew",sep="\t")

def do(self,action):
    """updates the state based on the agent doing action. returns state,reward """
    reward = 0.0
    # A prize can appear:
    if self.prize is None and flip(self.prize_appears_prob):
        self.prize = random.choice(self.prize_locs)
    # Actions can be noisy
    if flip(0.4):
        actual_direction = random.choice(self.actions)
    else:
        actual_direction = action
    # Modeling the actions given the actual direction
    if actual_direction == "right":
        if self.x==self.xdim-1 or (self.x, self.y) in self.vwalls:
            reward += self.crashed_reward
        else:
            self.x += 1
    elif actual_direction == "left":
        if self.x==0 or (self.x-1, self.y) in self.vwalls:
            reward += self.crashed_reward
else:
    self.x += -1
elif actual_direction == "up":
    if self.y==self.ydim-1:
        reward += self.crashed_reward
    else:
        self.y += 1
elif actual_direction == "down":
    if self.y==0:
        reward += self.crashed_reward
    else:
        self.y += -1
else:
    raise RuntimeError("unknown_direction"+str(direction))

# Monsters
if (self.x,self.y) in self.monster_locs and flip(self.monster_appears_prob):
    if self.damaged:
        reward += self.monster_reward_when_damaged
    else:
        self.damaged = True
if (self.x,self.y) in self.repair_stations:
    self.damaged = False

# Prizes
if (self.x,self.y) == self.prize:
    reward += self.prize_reward
    self.prize = None

# Statistics
self.number_steps += 1
self.total_reward += reward
if self.total_reward < self.min_reward:
    self.min_reward = self.total_reward
    self.min_step = self.number_steps
if self.total_reward>0 and reward>self.total_reward:
    self.zero_crossing = self.number_steps
self.display(2,"",self.number_steps,self.total_reward,
    self.total_reward/self.number_steps,sep="\t")

return (self.x, self.y, self.damaged, self.prize), reward

13.1.3 Evaluation and Plotting

import matplotlib.pyplot as plt

def plot_rl(ag, label=None, yplot='Total', step_size=None,

http://aipython.org  Version 0.9.4  March 2, 2022
13. Reinforcement Learning

```python
    steps_explore=1000, steps_exploit=1000, xscale='linear'):
        """
        plots the agent ag
        label is the label for the plot
        yplot is 'Average' or 'Total'
        step_size is the number of steps between each point plotted
        steps_explore is the number of steps the agent spends exploring
        steps_exploit is the number of steps the agent spends exploiting
        xscale is 'log' or 'linear'
        
        returns total reward when exploring, total reward when exploiting
        """
        assert yplot in ['Average','Total']
        if step_size is None:
            step_size = max(1,(steps_explore+steps_exploit)//500)
        if label is None:
            label = ag.label
        ag.max_display_level,old_mdl = 1,ag.max_display_level
        plt.ion()
        plt.xscale(xscale)
        plt.xlabel("step")
        plt.ylabel(yplot+" reward")
        steps = []  # steps
        rewards = []  # return
        ag.restart()
        step = 0
        while step < steps_explore:
            ag.do(step_size)
            step += step_size
            steps.append(step)
            if yplot == "Average":
                rewards.append(ag.acc_rewards/step)
            else:
                rewards.append(ag.acc_rewards)
        acc_rewards_exploring = ag.acc_rewards
        ag.explore,explore_save = 0,ag.explore
        while step < steps_explore+steps_exploit:
            ag.do(step_size)
            step += step_size
            steps.append(step)
            if yplot == "Average":
                rewards.append(ag.acc_rewards/step)
            else:
                rewards.append(ag.acc_rewards)
        plt.plot(steps,rewards,label=label)
        plt.legend(loc="upper left")
        plt.draw()
        ag.max_display_level = old_mdl
        ag.explore=explore_save
        return acc_rewards_exploring, ag.acc_rewards-acc_rewards_exploring
```

http://aipython.org
Version 0.9.4
March 2, 2022
13.2 Q Learning

To run the Q-learning demo, in folder “aipython”, load “rlQTest.py”, and copy and paste the example queries at the bottom of that file. This assumes Python 3.

```python
import random
from display import Displayable
from utilities import argmaxe, flip

class RL_agent(Displayable):
    """An RL_Agent has percepts (s, r) for some state s and real reward r
    ""

class Q_learner(RL_agent):
    """A Q-learning agent has belief-state consisting of
    state is the previous state
    q is a {(state,action):value} dict
    visits is a {(state,action):n} dict. n is how many times action was
done in state
    acc_rewards is the accumulated reward
    it observes (s, r) for some world-state s and real reward r
    ""

def __init__(self, env, discount, explore=0.1, fixed_alpha=True,
            alpha=0.2,
            alpha_fun=lambda k:1/k,
            qinit=0, label="Q_learner"):
    """env is the environment to interact with.
discount is the discount factor
explore is the proportion of time the agent will explore
fixed_alpha specifies whether alpha is fixed or varies with the
number of visits
alpha is the weight of new experiences compared to old experiences
alpha_fun is a function that computes alpha from the number of
visits
qinit is the initial value of the Q's
label is the label for plotting
""
    RL_agent.__init__(self)
    self.env = env
    self.actions = env.actions
```

http://aipython.org

Version 0.9.4 March 2, 2022
self.discount = discount
self.explore = explore
self.fixed_alpha = fixed_alpha
self.alpha = alpha
self.alpha_fun = alpha_fun
self.qinit = qinit
self.label = label
self.restart()

restart is used to make the learner relearn everything. This is used by the plotter to create new plots.

def restart(self):
    """make the agent relearn, and reset the accumulated rewards
    """
    self.acc_rewards = 0
    self.state = self.env.state
    self.q = {}
    self.visits = {}  

do takes in the number of steps.

def do(self,num_steps=100):
    """do num_steps of interaction with the environment"
    self.display(2,"s\ta\tr\ts\tQ")
    alpha = self.alpha
    for i in range(num_steps):
        action = self.select_action(self.state)
        next_state,reward = self.env.do(action)
        if not self.fixed_alpha:
            k = self.visits[(self.state, action)] =
                 self.visits.get((self.state, action),0)+1
            alpha = self.alpha_fun(k)
            self.q[(self.state, action)] =
                 ((1-alpha) * self.q.get((self.state, action),self.qinit) +
                 alpha * (reward + self.discount * 
                 max(self.q.get((next_state,
                 next_act),self.qinit))
                 for next_act in self.actions)))
        self.display(2,self.state, action, reward, next_state,
                 self.q[(self.state, action)], sep='t')
        self.state = next_state
        self.acc_rewards += reward

select_action is used to select the next action to perform. This can be reimplemented to give a different exploration strategy.
13.2. Q Learning

Given the state, and the q-function

```python
if flip(self.explore):
    return random.choice(self.actions)
else:
    return argmaxe((next_act, self.q.get((state, next_act), self.qinit))
                      for next_act in self.actions)
```

Exercise 13.1  Implement a soft-max action selection. Choose a temperature that works well for the domain. Explain how you picked this temperature. Compare the epsilon-greedy, soft-max and optimism in the face of uncertainty.

Exercise 13.2  Implement SARSA. Hint: it does not do a max in do. Instead it needs to choose `next_act` before it does the update.

13.2.1 Testing Q-learning

The first tests are for the 2-action 2-state

```python
from rlProblem import Healthy_env
from rLOrner import Q_learner
from rLPlot import plot_rl

env = Healthy_env()
ad = Q_learner(env, 0.7)
ad_opt = Q_learner(env, 0.7, qinit=100, label="optimistic") # optimistic agent
ad_exp_l = Q_learner(env, 0.7, explore=0.01, label="less explore")
ad_exp_m = Q_learner(env, 0.7, explore=0.5, label="more explore")
ad_disc = Q_learner(env, 0.9, qinit=100, label="disc 0.9")
ad_va = Q_learner(env, 0.7, qinit=100, fixed_alpha=False, alpha_fun=lambda k:10/(9+k), label="alpha=10/(9+k)"

# ad.max_display_level = 2
# ad.do(20)
# ad.q  # get the learned q-values
# ad.max_display_level = 1
# ad.do(1000)
# ad.q  # get the learned q-values
# plot_rl(ad,yplot="Average")
# plot_rl(ad_opt,yplot="Average")
# plot_rl(ad_exp_l,yplot="Average")
# plot_rl(ad_exp_m,yplot="Average")
# plot_rl(ad_disc,yplot="Average")
# plot_rl(ad_va,yplot="Average")

from mdpExamples import MDPtiny
from rlProblem import Env_from_MDP
envt = Env_from_MDP(MDPtiny())
```

http://aipython.org
13. Reinforcement Learning

```python
agt = Q_learner(envt, 0.8)
# agt.do(20)

from rlSimpleEnv import Simple_game_env
senv = Simple_game_env()
sag1 = Q_learner(senv, 0.9, explore=0.2, fixed_alpha=True, alpha=0.1)
#
# plot_rl(sag1, steps_explore=100000, steps_exploit=100000, label="alpha=\"+str(sag1.alpha))

sg2 = Q_learner(senv, 0.9, explore=0.2, fixed_alpha=False)
# plot_rl(sg2, steps_explore=100000, steps_exploit=100000, label="alpha=1/k")

sag3 = Q_learner(senv, 0.9, explore=0.2, fixed_alpha=False, alpha_fun=lambda k: 10/(9+k))
#
# plot_rl(sag3, steps_explore=100000, steps_exploit=100000, label="alpha=10/(9+k")
```

### 13.3 Q-learning with Experience Replay

Warning: not properly debugged

```python
from rlQLearner import Q_learner
from utilities import flip
import random

class BoundedBuffer(object):
    def __init__(self, buffer_size=1000):
        self.buffer_size = buffer_size
        self.buffer = [0]*buffer_size
        self.number_added = 0

    def add(self, experience):
        if self.number_added < self.buffer_size:
            self.buffer[self.number_added] = experience
        else:
            if flip(self.buffer_size/self.number_added):
                position = random.randrange(self.buffer_size)
                self.buffer[position] = experience
            self.number_added += 1

    def get(self):
        return self.buffer[random.randrange(min(self.number_added, self.buffer_size))]

class Q_AR_learner(Q_learner):
    def __init__(self, env, discount, explore=0.1, fixed_alpha=True, alpha=0.2,
                 alpha_fun=lambda k: 1/k, qinit=0, label="Q_AR_learner",
                 max_buffer_size=5000, num_updates_per_action=5, burn_in=1000):
```

[http://aipython.org](http://aipython.org)
Q_learner.__init__(self, env, discount, explore, fixed_alpha, alpha,
alpha_fun, qinit, label)
self.experience_buffer = BoundedBuffer(max_buffer_size)
self.num_updates_per_action = num_updates_per_action
self.burn_in = burn_in

def do(self,num_steps=100):
    """do num_steps of interaction with the environment""
    self.display(2,"s\ta\tr\ts\tQ")
    alpha = self.alpha
    for i in range(num_steps):
        action = self.select_action(self.state)
        next_state,reward = self.env.do(action)
        self.experience_buffer.add((self.state,action,reward,next_state))
        #remember experience
        if not self.fixed_alpha:
            k = self.visits[(self.state, action)] =
            self.visits.get((self.state, action),0)+1
            alpha = self.alpha_fun(k)
            self.q[(self.state, action)] = (1-alpha) * self.q.get((self.state, action),self.qinit)
            + alpha * (reward + self.discount
                           * max(self.q.get((next_state,
                                           next_act),self.qinit)
                           for next_act in self.actions))
        self.display(2,self.state, action, reward, next_state,
                      self.q[(self.state, action)], sep='\t')
        self.state = next_state
        self.acc_rewards += reward
        # do some updates from experince buffer
        if self.experience_buffer.number_added > self.burn_in:
            for i in range(self.num_updates_per_action):
                (s,a,r,ns) = self.experience_buffer.get()
                if not self.fixed_alpha:
                    k = self.visits[(s,a)]
                    alpha = self.alpha_fun(k)
                    self.q[(s,a)] = (1-alpha) * self.q[(s,a)]
                    + alpha * (reward + self.discount
                                 * max(self.q.get((ns,na),self.qinit)
                                 for na in self.actions))

from rlSimpleEnv import Simple_game_env
from rlQTest import sag1, sag2, sag3
from rlPlot import plot_rl

senv = Simple_game_env()
saglar = Q_AR_learner(senv,0.9,explore=0.2,fixed_alpha=True,alpha=0.1)
# Reinforcement Learning

To run the demo, in folder "aipython", load "rlModelLearner.py", and copy and paste the example queries at the bottom of that file. This assumes Python 3.

A model-based reinforcement learner builds a Markov decision process model of the domain, simultaneously learns the model and plans with that model.

The model-based reinforcement learner used the following data structures:

- \( q[s,a] \) is dictionary that, given a \( (s,a) \) pair returns the Q-value, the estimate of the future (discounted) value of being in state \( s \) and doing action \( a \).

- \( r[s,a] \) is dictionary that, given a \( (s,a) \) pair returns the average reward from doing \( a \) in state \( s \).

- \( t[s,a,s'] \) is dictionary that, given a \( (s,a,s') \) tuple returns the number of times \( a \) was done in state \( s \), with the result being state \( s' \).

- \( visits[s,a] \) is dictionary that, given a \( (s,a) \) pair returns the number of times action \( a \) was carried out in state \( s \).

- \( res\_states[s,a] \) is dictionary that, given a \( (s,a) \) pair returns the list of resulting states that have occurred when action \( a \) was carried out in state \( s \). This is used in the asynchronous value iteration to determine the \( s' \) states to sum over.

- \( visits\_list \) is a list of \( (s,a) \) pair that have been carried out. This is used to ensure there is no divide-by zero in the asynchronous value iteration. Note that this could be constructed from \( r, visits \) or \( res\_states \) by enumerating the keys, but needs to be a list for \texttt{random.choice}, and we don’t want to keep recreating it.

http://aipython.org  
Version 0.9.4  
March 2, 2022
import random
from rQLearner import RL_agent
from display import Displayable
from utilities import argmaxe, flip

class Model_based_reinforcement_learner(RL_agent):
    """A Model-based reinforcement learner
    ""
    def __init__(self, env, discount, explore=0.1, qinit=0,
                 updates_per_step=10, label="MBR_learner"):
        """env is the environment to interact with.
        discount is the discount factor
        explore is the proportion of time the agent will explore
        qinit is the initial value of the Q's
        updates_per_step is the number of AVI updates per action
        label is the label for plotting
        ""
        RL_agent.__init__(self)
        self.env = env
        self.actions = env.actions
        self.discount = discount
        self.explore = explore
        self.qinit = qinit
        self.updates_per_step = updates_per_step
        self.label = label
        self.restart()

    def restart(self):
        """make the agent relearn, and reset the accumulated rewards
        ""
        self.acc_rewards = 0
        self.state = self.env.state
        self.q = {}  # {(st,action):q_value} map
        self.r = {}  # {(st,action):reward} map
        self.t = {}  # {(st,action,st_next):count} map
        self.visits = {}  # {(st,action):count} map
        self.res_states = {}  # {(st,action):set_of_states} map
        self.visits_list = []  # list of (st,action)
        self.previous_action = None

    def do(self,num_steps=100):
        """do num_steps of interaction with the environment
        for each action, do updates_per_step iterations of asynchronous
        value iteration
        ""
        for step in range(num_steps):
            # do action, get reward, and update Q-values
http://aipython.org  Version 0.9.4  March 2, 2022
pst = self.state  # previous state
action = self.select_action(pst)
self.state, reward = self.env.do(action)
self.acc_rewards += reward
self.t[(pst, action, self.state)] = self.t.get((pst, action, self.state), 0)+1
if (pst, action) in self.visits:
    self.visits[(pst, action)] += 1
    self.r[(pst, action)] +=
        (reward-self.r[(pst, action)])/self.visits[(pst, action)]
    self.res_states[(pst, action)].add(self.state)
else:
    self.visits[(pst, action)] = 1
    self.r[(pst, action)] = reward
    self.res_states[(pst, action)] = {self.state}
    self.visits_list.append((pst, action))

st, act = pst, action  # initial state-action pair for AVI
for update in range(self.updates_per_step):
    self.q[(st, act)] = self.r[(st, act)]+self.discount*
        sum(self.t[(st, act, rst)]/self.visits[(st, act)]*
            max(self.q.get((rst, nact), self.qinit) for nact in self.actions)
            for rst in self.res_states[(st, act)])
    st, act = random.choice(self.visits_list)

---_rlModelLearner.py — (continued)______________________________

```python
def select_action(self, state):
    """returns an action to carry out for the current agent
    given the state, and the q-function
    ""
    if flip(self.explore):
        return random.choice(self.actions)
    else:
        return argmaxe((next_act, self.q.get((state, next_act), self.qinit))
                        for next_act in self.actions)
```

---_rlModelLearner.py — (continued)______________________________

```python
from rlQTest import senv  # simple game environment
mbl1 = Model_based_reinforcement_learner(senv, 0.9, updates_per_step=10)
    # plot_rl(mbl1, steps_explore=100000, steps_exploit=100000, label="model-based(10)")
mbl2 = Model_based_reinforcement_learner(senv, 0.9, updates_per_step=1)
    # plot_rl(mbl2, steps_explore=100000, steps_exploit=100000, label="model-based(1)"

Exercise 13.3 If there was only one update per step, the algorithm can be made
simpler and use less space. Explain how. Does it make it more efficient? Is it
worthwhile having more than one update per step for the games implemented
here?
```

http://aipython.org  Version 0.9.4  March 2, 2022
Exercise 13.4  It is possible to implement the model-based reinforcement learner by replacing \( q, r, \) \( \text{visits}, \text{res states} \) with a single dictionary that returns a tuple \((q,r,v,tm)\) where \( q, r \) and \( v \) are numbers, and \( tm \) is a map from resulting states into counts. Does this make the algorithm easier to understand? Does this make the algorithm more efficient?

Exercise 13.5  If the states and the actions were mapped into integers, the dictionaries could be implemented more efficiently as arrays. This entails an extra step in specifying problems. Implement this for the simple game. Is it more efficient?

13.5  Reinforcement Learning with Features

To run the demo, in folder “aipython”, load “rlFeatures.py”, and copy and paste the example queries at the bottom of that file. This assumes Python 3.

13.5.1  Representing Features

A feature is a function from state and action. To construct the features for a domain, we construct a function that takes a state and an action and returns the list of all feature values for that state and action. This feature set is redesigned for each problem.

\[ \text{get_features(state, action)} \]
returns the feature values appropriate for the simple game.

```python
from rlSimpleEnv import Simple_game_env
from rlProblem import RL_env

def get_features(state, action):
    """returns the list of feature values for the state-action pair
    """
    assert action in Simple_game_env.actions
    (x, y, d, p) = state
    # f1: would go to a monster
    f1 = monster_ahead(x, y, action)
    # f2: would crash into wall
    f2 = wall_ahead(x, y, action)
    # f3: action is towards a prize
    f3 = towards_prize(x, y, action, p)
    # f4: damaged and action is toward repair station
    f4 = towards_repair(x, y, action) if d else 0
    # f5: damaged and towards monster
    f5 = 1 if d and f1 else 0
    # f6: damaged
    f6 = 1 if d else 0
    # f7: not damaged
```

http://aipython.org  Version 0.9.4  March 2, 2022
f7 = 1 - f6
# f8: damaged and prize ahead
f8 = 1 if d and f3 else 0
# f9: not damaged and prize ahead
f9 = 1 if not d and f3 else 0
features = [1, f1, f2, f3, f4, f5, f6, f7, f8, f9]
# the next 20 features are for 5 prize locations
# and 4 distances from outside in all directions
for pr in Simple_game_env.prize_locs+[None]:
    if p==pr:
        features += [x, 4-x, y, 4-y]
    else:
        features += [0, 0, 0, 0]
# fp04 feature for y when prize is at 0,4
# this knows about the wall to the right of the prize
if p==(0,4):
    if x==0:
        fp04 = y
    elif y<3:
        fp04 = y
    else:
        fp04 = 4-y
else:
    fp04 = 0
features.append(fp04)
return features

def monster_ahead(x,y,action):
    """returns 1 if the location expected to get to by doing
    action from (x,y) can contain a monster.
    """
    if action == "right" and (x+1,y) in Simple_game_env.monster_locs:
        return 1
    elif action == "left" and (x-1,y) in Simple_game_env.monster_locs:
        return 1
    elif action == "up" and (x,y+1) in Simple_game_env.monster_locs:
        return 1
    elif action == "down" and (x,y-1) in Simple_game_env.monster_locs:
        return 1
    else:
        return 0

def wall_ahead(x,y,action):
    """returns 1 if there is a wall in the direction of action from (x,y).
    This is complicated by the internal walls.
    """
    if action == "right" and (x==Simple_game_env.xdim-1 or (x,y) in
        Simple_game_env.vwalls):
        return 1
    elif action == "left" and (x==0 or (x-1,y) in Simple_game_env.vwalls):
    
http://aipython.org
Version 0.9.4
March 2, 2022
elif action == "up" and y==Simple_game_env.ydim-1:
    return 1
eelif action == "down" and y==0:
    return 1
eelse:
    return 0
def towards_prize(x,y,action,p):
    """action goes in the direction of the prize from (x,y)""
    if p is None:
        return 0
eelif p==(0,4): # take into account the wall near the top-left prize
        if action == "left" and (x>1 or x==1 and y<3):
            return 1
eelif action == "down" and (x>0 and y>2):
            return 1
eelif action == "up" and (x==0 or y<2):
            return 1
eelse:
return 0
eelse:
    px,py = p
    if p==(4,4) and x==0:
        if (action=="right" and y<3) or (action=="down" and y<3) or (action=="up" and y<2):
            return 1
eelse:
return 0
eelif action == "up" and y<py) or (action == "down" and py<y):
return 1
eelif (action == "left" and px<x) or (action == "right" and x<px):
return 1
eelse:
return 0
def towards_repair(x,y,action):
    """returns 1 if action is towards the repair station.
    ""
    if action == "up" and (x>0 and y<4 or x==0 and y<2):
        return 1
eelif action == "left" and x>1:
        return 1
eelif action == "right" and x==0 and y<3:
        return 1
eelif action == "down" and x==0 and y>2:
        return 1
eelse:
return 0
def simp_features(state, action):
    
    # returns a list of feature values for the state-action pair
    assert action in Simple_game_env.actions
    (x, y, d, p) = state
    # f1: would go to a monster
    f1 = monster_ahead(x, y, action)
    # f2: would crash into wall
    f2 = wall_ahead(x, y, action)
    # f3: action is towards a prize
    f3 = towards_prize(x, y, action, p)
    return [1, f1, f2, f3]

13.5.2 Feature-based RL learner

This learns a linear function approximation of the Q-values. It requires the
gradient descent step size

13.5. Reinforcement Learning with Features

```python
self.get_features = get_features
self.actions = env.actions
self.discount = discount
self.explore = explore
self.step_size = step_size
self.winit = winit
self.label = label
self.restart()
```

`restart()` is used to make the learner relearn everything. This is used by the plotter to create new plots.

```python
def restart(self):
    """make the agent relearn, and reset the accumulated rewards
    """
    self.acc_rewards = 0
    self.state = self.env.state
    self.features = self.get_features(self.state,
        list(self.env.actions)[0])
    self.weights = [self.winit for f in self.features]
    self.action = self.select_action(self.state)
```

`do` takes in the number of steps.

```python
def do(self, num_steps=100):
    """do num_steps of interaction with the environment"
    self.display(2, "s\ta\tr\ts'\tQ\tdelta")
    for i in range(num_steps):
        next_state, reward = self.env.do(self.action)
        self.acc_rewards += reward
        next_action = self.select_action(next_state)
        feature_values = self.get_features(self.state, self.action)
        self.display(2, self.state, self.action, reward, next_state,
            dot_product(self.weights, feature_values), delta,
            sep='\t')
        self.state = next_state
        self.action = next_action
```

```python
def select_action(self, state):
    """returns an action to carry out for the current agent
    given the state, and the q-function.
    This implements an epsilon-greedy approach
    where self.explore is the probability of exploring.
    """
```
if flip(self.explore):
    return random.choice(self.actions)
else:
    return argmaxe((next_act, dot_product(self.weights,
        self.get_features(state,next_act)))
        for next_act in self.actions)

def show_actions(self,state=None):
    """prints the value for each action in a state.
    This may be useful for debugging.
    """
    if state is None:
        state = self.state
    for next_act in self.actions:
        print(next_act,dot_product(self.weights,
            self.get_features(state,next_act)))

def dot_product(l1,l2):
    return sum(e1*e2
        for (e1,e2) in zip(l1,l2))

Test code:

from rlQTest import * # simple game environment
from rlSimpleGameFeatures import get_features, simp_features
from rlPlot import *

fa1 = SARSA_LFA_learner(senv, get_features, 0.9, step_size=0.01)
    #fa1.max_display_level = 2
    #fa1.do(20)
    #plot_rl(fa1,steps_explore=10000,steps_exploit=10000,label="SARSA_LFA(0.01)"
fas1 = SARSA_LFA_learner(senv, simp_features, 0.9, step_size=0.01)
    #plot_rl(fas1,steps_explore=10000,steps_exploit=10000,label="SARSA_LFA(simp)"

Exercise 13.6 How does the step-size affect performance? Try different step sizes (e.g., 0.1, 0.001, other sizes in between). Explain the behaviour you observe. Which step size works best for this example. Explain what evidence you are basing your prediction on.


Exercise 13.8 For each of the following first predict, then plot, then explain the behaviour you observed:

(a) SARSA_LFA, Model-based learning (with 1 update per step) and Q-learning for 10,000 steps 20% exploring followed by 10,000 steps 100% exploiting

(b) SARSA_LFA, model-based learning and Q-learning for

i) 100,000 steps 20% exploring followed by 100,000 steps 100% exploit
ii) 10,000 steps 20% exploring followed by 190,000 steps 100% exploit

http://aipython.org
13.5. Reinforcement Learning with Features

(c) Suppose your goal was to have the best accumulated reward after 200,000 steps. You are allowed to change the exploration rate at a fixed number of steps. For each of the methods, which is the best position to start exploiting more? Which method is better? What if you wanted to have the best reward after 10,000 or 1,000 steps?

Based on this evidence, explain when it is preferable to use SARSA_LFA, Model-based learner, or Q-learning.

Important: you need to run each algorithm more than once. Your explanation should include the variability as well as the typical behavior.

13.5.3 Experience Replay

Here we consider experience replay with a bounded replay buffer for SARSA_LFA. Warning: does not work properly yet.

Should self.env return (reward,state) to be consistent with (S,A,R,S)?

```python
from rlFeatures import SARSA_LFA_learner, dot_product
from utilities import flip
import random

class SARSA_LFA_AR_learner(SARSA_LFA_learner):
    def __init__(self, env, get_features, discount, explore=0.2, step_size=0.01, winit=0, label="SARSA_LFA-AR", max_buffer_size=500, num_updates_per_action=5, burn_in=100):
        SARSA_LFA_learner.__init__(self, env, get_features, discount, explore, step_size, winit, label)
        self.max_buffer_size = max_buffer_size
        self.action_buffer = [0]*max_buffer_size
        self.number_added = 0
        self.num_updates_per_action = num_updates_per_action
        self.burn_in = burn_in

    def add_to_buffer(self,experience):
        if self.number_added < self.max_buffer_size:
            self.action_buffer[self.number_added] = experience
        else:
            if flip(self.max_buffer_size/self.number_added):
                position = random.randrange(self.max_buffer_size)
                self.action_buffer[position] = experience
            self.number_added += 1

    def do(self, num_steps=100):
        """do num_steps of interaction with the environment""
        self.display(2, "S\ta\treward\tQ\tdelta")
        for i in range(num_steps):
```

http://aipython.org  Version 0.9.4  March 2, 2022
next_state, reward = self.env.do(self.action)
self.add_to_buffer((self.state, self.action, reward, next_state))
    # remember experience
self.acc_rewards += reward
next_action = self.select_action(next_state)
feature_values = self.get_features(self.state, self.action)
oldQ = dot_product(self.weights, feature_values)
nextQ = dot_product(self.weights, self.get_features(next_state, next_action))
delta = reward + self.discount * nextQ - oldQ
for i in range(len(self.weights)):
    self.weights[i] += self.step_size * delta * feature_values[i]
self.state = next_state
self.action = next_action
if self.number_added > self.burn_in:
    for i in range(self.num_updates_per_action):
        (s, a, r, ns) =
            self.action_buffer[random.randrange(min(self.number_added, self.max_buffer_size))]
        na = self.select_action(ns)
        feature_values = self.get_features(s, a)
        oldQ = dot_product(self.weights, feature_values)
        nextQ = dot_product(self.weights, self.get_features(ns, na))
        delta = reward + self.discount * nextQ - oldQ
        for i in range(len(self.weights)):
            self.weights[i] += self.step_size * delta * feature_values[i]

Test code:

from rlQTest import env # simple game environment
from rlSimpleGameFeatures import get_features, simp_features
from rlPlot import plot_rl
fa1 = SARSA_LFA_AR_learner(env, get_features, 0.9, step_size=0.01)
# fa1.max_display_level = 2
# fa1.do(20)
# plot_rl(fa1, steps_explore=10000, steps_exploit=10000, label="SARSA_LFA_AR(0.01)"
fas1 = SARSA_LFA_AR_learner(env, simp_features, 0.9, step_size=0.01)
# plot_rl(fas1, steps_explore=10000, steps_exploit=10000, label="SARSA_LFA_AR(simp)"

13.6 Multiagent Learning

The next code of for multiple agents that learn when interacting with other agents. This code is designed to be extended, and as such is restricted to being

http://aipython.org
two agents, a single state, and the only observation is the reward. Coordinating
agents can’t easily implement that agent architecture. However, in that archi-
tecture, an agent calls the environment. That architecture was chosen because
it was simple. However, it does not really work when there are multiple agents,
instead we have a controller that tells the agents the percepts (here the percepts
are just the reward).

```python
from display import Displayable
import utilities # argmaxall for (element,value) pairs
import matplotlib.pyplot as plt
import random

class GameAgent(Displayable):
    next_id=0
    def __init__(self, actions):
        
        Actions is the set of actions the agent can do. It needs to be told
        that!
        
        self.actions = actions
        self.id = GameAgent.next_id
        GameAgent.next_id += 1
        self.display(2,f"Agent {self.id} has actions {actions}"
        self.dist = {act:1 for act in actions} # unnormalized distibution
        self.total_score = 0

    def init_action(self):
        
        Act randomly initially
        Could be overridden (but I'm not sure why you would).
        
        self.act = random.choice(self.actions)
        return self.act

    def select_action(self, reward):
        
        Select the action given the reward.
        This implements "Act randomly" and should be overridden!
        
        self.total_score += reward
        self.act = random.choice(self.actions)
        return self.act

class SimpleCountingAgent(GameAgent):
    
    """This agent just counts the number of times (it thinks) it has won
    and does the
    actions it thinks is most likely to win.
    """

http://aipython.org
```
def __init__(self, actions, prior_count=1):
    """
    Actions is the set of actions the agent can do. It needs to be told
    that!
    """
    GameAgent.__init__(self, actions)
    self.prior_count = prior_count
    self.dist = {a: prior_count for a in self.actions} # unnormalized
distribution
    self.averew = 0
    self.num_steps = 0

def select_action(self, reward):
    self.total_score += reward
    self.num_steps += 1
    self.display(2,"The reward for agent {self.id} was {reward}"")
    self.averew = self.averew+(reward-self.averew)/self.num_steps
    if reward>self.averew:
        self.dist[self.act] += 1
    else:
        for otheract in self.actions:
            if otheract != self.act:
                self.dist[otheract] += 1/len(self.actions)
    self.display(2,"Distribution for agent {self.id} is 
(normalize(self.dist))")
    self.act = select_from_dist(self.dist)
    self.display(2,"Agent {self.id} did {self.act}"")
    return self.act

class SimpleQAgent(GameAgent):
    """This agent maintains the Q-function for each state.
(Or just the average reward as the future state is all the same).
Chooses the best action using
"""
    def __init__(self, actions, q_init=100, alpha=0.1,
                 prob_step_size=0.001, min_prob=0.01):
        """
        Actions is the set of actions the agent can do. It needs to be told
        that!
        q_init is the initial q-values
        alpha is the step size for action estimate
        prob_step_size is the step size for probability change
        min_prob is the minimum a probability should become
        """
        GameAgent.__init__(self, actions)
        self.Q = {a:q_init for a in self.actions}
        self.dist = normalize({a:0.7+random.random() for a in
                                 self.actions}) # start with random dist but not too close to
                                 zero

http://aipython.org  Version 0.9.4  March 2, 2022
13.6. Multiagent Learning

```python
self.alpha = alpha
self.prob_step_size = prob_step_size
self.min_prob = min_prob
self.num_steps = 1 # (1 because it is only used after initial step)

def select_action(self, reward):
    self.total_score += reward
    self.display(2, f"The reward for agent {self.id} was {reward}"
    a_best = utilities.argmaxall(self.Q.items())
    for a in self.actions:
        if a in a_best:
            self.dist[a] += self.prob_step_size
        else:
            self.dist[a] -= min(self.dist[a], self.prob_step_size)
            self.dist[a] = max(self.dist[a], self.min_prob)
    self.dist = normalize(self.dist)
    self.display(2, f"Distribution for agent {self.id} is {self.dist}"
    self.act = select_from_dist(self.dist)
    self.display(2, f"Agent {self.id} did {self.act}"
    return self.act

def normalize(dist):
    """unnorm dict is a {value:number} dictionary, where the numbers are
    all non-negative
    returns dict where the numbers sum to one"
    tot = sum(dist.values())
    return {var:val/tot for (var,val) in dist.items()}

def select_from_dist(dist):
    rand = random.random()
    for (act,prob) in normalize(dist).items():
        rand -= prob
        if rand < 0:
            return act
```

The simulator takes a game and simulates the game:

```python
class SimulateGame(Displayable):
    def __init__(self, game, agents):
        self.game = game
        self.agents = agents # list of agents
        self.action_history = []
        self.reward_history = []
        self.dist_history = []
        self.actions = tuple(ag.init_action() for ag in self.agents)
        self.num_steps = 0

    def go(self, steps):
```

http://aipython.org
```
for i in range(steps):
    self.num_steps += 1
    self.rewards = self.game.play(self.actions)
    self.reward_history.append(self.rewards)
    self.actions =
        tuple(self.agents[i].select_action(self.rewards[i])
        for i in range(self.game.num_agents))
    self.action_history.append(self.actions)
    self.dist_history.append([normalize(ag.dist)
                  for ag in self.agents])
    print("Scores:", ', '.join(f"Agent {ag.id} average reward={ag.total_score/self.num_steps}" for ag in self.agents))
#return self.reward_history, self.action_history

def action_dist(self, which_actions=[1,1]):
    """ which actions is [a0,a1]
    returns the empirical distribution of actions for agents,
    where ai specifies the index of the actions for agent i
    """
    return [sum(1 for a in sim.action_history
                  if a[i]==gm.actions[i][which_actions[i]])/len(sim.action_history)
               for i in range(2)]

def plot_dynamics(self, x_action=0, y_action=0):
    # make it interactive
    plt.ion()
    agents = self.agents
    x_act = self.game.actions[0][x_action]
    y_act = self.game.actions[1][y_action]
    plt.xlabel(f"Action {self.agents[0].actions[x_action]} for Agent {agents[0].id}")
    plt.ylabel(f"Action {self.agents[1].actions[y_action]} for Agent {agents[1].id}")
    plt.plot([self.dist_history[t][0][x_act] for t in range(len(self.dist_history))],
              [self.dist_history[t][1][y_act] for t in range(len(self.dist_history))],
            #plt.legend()

The following are some games from Poole and Mackworth [2017].
13.6. Multiagent Learning

```python
('football', 'shopping'): (0, 0),
('shopping', 'football'): (0, 0),
('shopping', 'shopping'): (1, 2)}[actions]

class SoccerGame(Displayable):
    def __init__(self):
        self.num_agents = 2
        self.actions = [['left', 'right']]*2

    def play(self, actions):
        return {('left', 'left'): (0.6, 0.4),
                ('left', 'right'): (0.2, 0.8),
                ('right', 'left'): (0.3, 0.7),
                ('right', 'right'): (0.9, 0.1)}[actions]

class GameShow(Displayable):
    def __init__(self):
        self.num_agents = 2
        self.actions = [['take', 'give']]*2

    def play(self, actions):
        return {('take', 'take'): (100, 100),
                ('take', 'give'): (1100, 0),
                ('give', 'take'): (0, 1100),
                ('give', 'give'): (1000, 1000)}[actions]

class UniqueNEGExample(Displayable):
    def __init__(self):
        self.num_agents = 2
        self.actions = [['a1', 'b1', 'c1'], ['d2', 'e2', 'f2']]

    def play(self, actions):
        return {('a1', 'd2'): (3, 5),
                ('a1', 'e2'): (5, 1),
                ('a1', 'f2'): (1, 2),
                ('b1', 'd2'): (1, 1),
                ('b1', 'e2'): (2, 9),
                ('b1', 'f2'): (6, 4),
                ('c1', 'd2'): (2, 6),
                ('c1', 'e2'): (4, 7),
                ('c1', 'f2'): (0, 8)}[actions]

# Choose one:
# gm = ShoppingGame()
# gm = SoccerGame()
```

http://aipython.org
# gm = GameShow()
# gm = UniqueNEGGameExample()

# Choose one:
# sim=SimulateGame(gm,[SimpleQAgent(gm.actions[0]),
    SimpleQAgent(gm.actions[1])]); sim.go(10000)
# sim= SimulateGame(gm,[SimpleCountingAgent(gm.actions[0]),
    SimpleCountingAgent(gm.actions[1])]); sim.go(10000)
# sim=SimulateGame(gm,[SimpleCountingAgent(gm.actions[0]),
    SimpleQAgent(gm.actions[1])]); sim.go(10000)

# sim.plot_dynamics()

# empirical proportion that agents did their action at index 1:
# sim.action_dist([1,1])

# learned distribution for agent 0
# sim.agents[0].dist
Chapter 14

Relational Learning

14.1 Collaborative Filtering


This assumes the form of the dataset from movielens (http://grouplens.org/datasets/movielens/). The rating are a set of \((user, item, rating, timestamp)\) tuples.

```python
import random
import matplotlib.pyplot as plt
import urllib.request
from learnProblem import Learner
from display import Displayable

class CF_learner(Learner):
    def __init__(self,
                  rating_set,  # a Rating_set object
                  rating_subset = None,  # subset of ratings to be used as training ratings
                  test_subset = None,  # subset of ratings to be used as test ratings
                  step_size = 0.01,  # gradient descent step size
                  reglz = 1.0,  # the weight for the regularization terms
                  num_properties = 10,  # number of hidden properties
                  property_range = 0.02 # properties are initialized to be between
                                           # -property_range and property_range
```

295
self.rating_set = rating_set
self.ratings = rating_subset or rating_set.training_ratings # whichever is not empty
if test_subset is None:
    self.test_ratings = self.rating_set.test_ratings
else:
    self.test_ratings = test_subset
self.step_size = step_size
self.reglz = reglz
self.num_properties = num_properties
self.num_ratings = len(self.ratings)
self.ave_rating = (sum(r for (u,i,r,t) in self.ratings) / self.num_ratings)
self.users = {u for (u,i,r,t) in self.ratings}
self.items = {i for (u,i,r,t) in self.ratings}
self.user_bias = {u:0 for u in self.users}
self.item_bias = {i:0 for i in self.items}
self.user_prop = {u:[random.uniform(-property_range,property_range) for p in range(num_properties)] for u in self.users}
self.item_prop = {i:[random.uniform(-property_range,property_range) for p in range(num_properties)] for i in self.items}
self.zeros = [0 for p in range(num_properties)]
self.iter=0

def stats(self):
    self.display(1, "ave sumsq error of mean for training=",
    sum((self.ave_rating-rating)**2 for (user,item,rating,timestamp) in self.ratings)/len(self.ratings))
    self.display(1, "ave sumsq error of mean for test=",
    sum((self.ave_rating-rating)**2 for (user,item,rating,timestamp) in self.test_ratings)/len(self.test_ratings))
    self.display(1, "error on training set",
    self.evaluate(self.ratings))
    self.display(1, "error on test set",
    self.evaluate(self.test_ratings))

learn carries out num_iter steps of gradient descent.

def prediction(self,user,item):
    """Returns prediction for this user on this item.
    The use of .get() is to handle users or items not in the training set.
    """
    return (self.ave_rating
          + self.user_bias.get(user,0) #self.user_bias[user]
14.1. Collaborative Filtering

```python
+ self.item_bias.get(item, 0) # self.item_bias[item]
+ 
+ sum([self.user_prop.get(user, self.zeros)[p] * self.item_prop.get(item, self.zeros)
+ for p in range(self.num_properties)])
```

```python
def learn(self, num_iter = 50):
    """do num_iter iterations of gradient descent."""
    for i in range(num_iter):
        self.iter += 1
        abs_error=0
        sumsq_error=0
        for (user, item, rating, timestamp) in random.sample(self.ratings, len(self.ratings)):
            error = self.prediction(user, item) - rating
            abs_error += abs(error)
            sumsq_error += error * error
            self.user_bias[user] -= self.step_size * error
            self.item_bias[item] -= self.step_size * error
            for p in range(self.num_properties):
                self.user_prop[user][p] -=
                self.step_size * error * self.item_prop[item][p]
                self.item_prop[item][p] -=
                self.step_size * error * self.user_prop[user][p]
        for user in self.users:
            self.user_bias[user] -= self.step_size * self.reglz *
            self.user_bias[user]
            for p in range(self.num_properties):
                self.user_prop[user][p] -=
                self.step_size * self.reglz * self.user_prop[user][p]
        for item in self.items:
            self.item_bias[item] -=
            self.step_size * self.reglz * self.item_bias[item]
            for p in range(self.num_properties):
                self.item_prop[item][p] -=
                self.step_size * self.reglz * self.item_prop[item][p]
        self.display(1, "Iteration", self.iter,
        "(Ave Abs, AveSumSq) training",
        "test =", self.evaluate(self.test_ratings))
```

`evaluate` evaluates current predictions on the rating set:

```python
def evaluate(self, ratings):
    """returns (average_absolute_error, average_sum_squares_error) for ratings"
    abs_error = 0
    sumsq_error = 0
    if not ratings: return (0, 0)
    for (user, item, rating, timestamp) in ratings:
```

http://aipython.org
Version 0.9.4
March 2, 2022
error = self.prediction(user, item) - rating
abs_error += abs(error)
sumsq_error += error * error
return abs_error/len(ratings), sumsq_error/len(ratings)

14.1.1 Alternative Formulation
An alternative formulation is to regularize after each update.

14.1.2 Plotting

```python
def plot_predictions(self, examples="test"):
    
    examples is either "test" or "training" or the actual examples
    
    if examples == "test":
        examples = self.test_ratings
    elif examples == "training":
        examples = self.ratings
    plt.ion()
    plt.xlabel("prediction")
    plt.ylabel("cumulative proportion")
    self.actuals = [[] for r in range(0,6)]
    for (user, item, rating, timestamp) in examples:
        self.actuals[rating].append(self.prediction(user, item))
    for rating in range(1,6):
        self.actuals[rating].sort()
        numrat = len(self.actuals[rating])
        yvals = [i/numrat for i in range(numrat)]
        plt.plot(self.actuals[rating], yvals,
        label="rating="+str(rating))
    plt.legend()
    plt.draw()
```

This plots a single property. Each (user, item, rating) is plotted where the x-value is the value of the property for the user, the y-value is the value of the property for the item, and the rating is plotted at this (x, y) position. That is, rating is plotted at the (x, y) position (p(user), p(item)).

```python
def plot_property(self, p, # property
    plot_all=False, # true if all points should be plotted
    num_points=200 # number of random points plotted if not all
    ):
    
    """plot some of the user-movie ratings,
```

http://aipython.org
Version 0.9.4
March 2, 2022
if plot_all is true
num_points is the number of points selected at random plotted.

the plot has the users on the x-axis sorted by their value on
property p and
with the items on the y-axis sorted by their value on property p and
the ratings plotted at the corresponding x-y position.

plt.ion()
plt.xlabel("users")
plt.ylabel("items")
user_vals = [self.user_prop[u][p]
    for u in self.users]
item_vals = [self.item_prop[i][p]
    for i in self.items]
plt.axis([min(user_vals)-0.02,
          max(user_vals)+0.05,
          min(item_vals)-0.02,
          max(item_vals)+0.05])
if plot_all:
    for (u,i,r,t) in self.ratings:
        plt.text(self.user_prop[u][p],
                 self.item_prop[i][p],
                 str(r))
else:
    for i in range(num_points):
        (u,i,r,t) = random.choice(self.ratings)
        plt.text(self.user_prop[u][p],
                 self.item_prop[i][p],
                 str(r))
plt.show()

14.1.3 Creating Rating Sets

A rating set can be read from the Internet or read from a local file. The default
is to read the Movielens 100K dataset from the Internet. It would be more
efficient to save the dataset as a local file, and then set local_file = True, as then
it will not need to download the dataset every time the program is run.

class Rating_set(Displayable):
def __init__(self,
    date_split=892000000,
    local_file=False,
    url="http://files.grouplens.org/datasets/movielens/ml-100k/u.data",
    file_name="u.data"):
    self.display(1,"reading...")
    if local_file:
        lines = open(file_name,'r')
    else:
        http://aipython.org
lines = (line.decode('utf-8') for line in urllib.request.urlopen(url))
all_ratings = (tuple(int(e) for e in line.strip().split('	'))
    for line in lines)
self.training_ratings = []
self.training_stats = {1:0, 2:0, 3:0, 4:0 ,5:0}
self.test_ratings = []
self.test_stats = {1:0, 2:0, 3:0, 4:0 ,5:0}
for rate in all_ratings:
        self.training_ratings.append(rate)
        self.training_stats[rate[2]] += 1
    else:
        self.test_ratings.append(rate)
        self.test_stats[rate[2]] += 1
self.display(1,...read:,, len(self.training_ratings),"trainingratings and",",
    len(self.test_ratings),"test ratings")
tr_users = {user for (user,item,rating,timestamp) in self.training_ratings}
test_users = {user for (user,item,rating,timestamp) in self.test_ratings}
selld.display(1,"users:","len(tr_users),"training","len(test_users),"test",",
    len(tr_users & test_users),"in common")
tr_items = {item for (user,item,rating,timestamp) in self.training_ratings}
test_items = {item for (user,item,rating,timestamp) in self.test_ratings}
selld.display(1,"items:","len(tr_items),"training","len(test_items),"test",",
    len(tr_items & test_items),"in common")
selld.display(1,"Rating statistics for training set:
    ",self.training_stats)
selld.display(1,"Rating statistics for test set: ",self.test_stats)

Sometimes it is useful to plot a property for all (user,item,rating) triples. There are too many such triples in the data set. The method create_top_subset creates a much smaller dataset where this makes sense. It picks the most rated items, then picks the users who have the most ratings on these items. It is designed for depicting the meaning of properties, and may not be useful for other purposes.

def create_top_subset(self, num_items = 30, num_users = 30):
    """Returns a subset of the ratings by picking the most rated items, and then the users that have most ratings on these, and then all of the ratings that involve these users and items."
    """
    items = {item for (user,item,rating,timestamp) in self.training_ratings}
14.1. Collaborative Filtering

```python
item_counts = {i:0 for i in items}
for (user,item,rating,timestamp) in self.training_ratings:
    item_counts[item] += 1

items_sorted = sorted((item_counts[i],i) for i in items)
top_items = items_sorted[-num_items:]
set_top_items = set(item for (count, item) in top_items)

users = {user for (user,item,rating,timestamp) in self.training_ratings}
user_counts = {u:0 for u in users}
for (user,item,rating,timestamp) in self.training_ratings:
    if item in set_top_items:
        user_counts[user] += 1

users_sorted = sorted((user_counts[u],u) for u in users)
top_users = users_sorted[-num_users:]
set_top_users = set(user for (count, user) in top_users)
used_ratings = [(user,item,rating,timestamp) for (user,item,rating,timestamp) in self.training_ratings]
    if user in set_top_users and item in set_top_items]
return used_ratings
```

movielens = Rating_set()
learner1 = CF_learner(movielens, num_properties = 1)
#learner1.learn(50)
# learner1.plot_predictions(examples = "training")
# learner1.plot_predictions(examples = "test")
#learner1.plot_property(0)
#movielens_subset = movielens.create_top_subset(num_items = 20, num_users = 20)
#learner_s = CF_learner(movielens, rating_subset=movielens_subset, test_subset=[], num_properties=1)
#learner_s.learn(1000)
#learner_s.plot_property(0,plot_all=True)
```
Chapter 15

Version History

• 2021-07-08 Version 0.9.1 updated the CSP code to have the same representation of variables as used by the probability code

• 2021-05-13 Version 0.9.0 Major revisions to chapters 8 and 9. Introduced recursive conditioning, simplified much code. New section on multiagent reinforcement learning.

• 2020-11-04 Version 0.8.6 simplified value iteration for MDPs.

• 2020-10-20 Version 0.8.4 planning simplified, and gives error if goal not part of state (by design). Fixed arc costs.

• 2020-07-21 Version 0.8.2 added positions and string to constraints

• 2019-09-17 Version 0.8.0 rerepresented blocks world (Section 6.1.2) due to bug found by Donato Meoli.
Bibliography


Index

\(\alpha-\beta\) pruning, 264
A* search, 41
A* Search, 44
action, 101
agent, 21, 267
argmax, 18
assignment, 52, 176
assumable, 97
asynchronous value iteration, 245
augmented feature, 136

batched stochastic gradient descent, 156
Bayesian network, 181
belief network, 181
blocks world, 104
Boolean feature, 126
bottom-up proof, 90
branch-and-bound search, 47
class
Action_instance, 118
Agent, 21
Arc, 34
Askable, 87
Assumable, 97

BNfromDBN, 220
BeliefNetwork, 181
Boosted_dataset, 157
Boosting_Learner, 157
Branch_and_bound, 84
CF_Learner, 295
CPDrename, 218
CSP, 53
CSP_from_STrips, 114
Clause, 87
Con_solver, 67
Constraint, 52
DBN, 218
DBNVEfilter, 220
DBNvariable, 217
DF_Branch_and_bound, 47
DT_Learner, 143
Data_from_file, 133
Data_from_files, 135
Data_set, 127
Data_set_augmented, 136
DecisionNetwork, 226
DecisionVariable, 225
Displayable, 17
EM_Learner, 253
Env_from_MDP, 268
| Environment | Evaluate | FactorDF | FactorMax | FactorObserved | FactorRename | FactorSum | Forward_STRIPS | FrontierPQ | GTB_learner | GibbsSampling | GraphicalModel | GridMDP | HMM | HMMVEfilter | HMM_Controlled | HMM_Local | HMMparticleFilter | Healthy,rv | InferenceMethod | KB | KBA | K_fold_dataset | K_means_learner | Layer | Learner | LikelihoodWeighting | Linear_complete_layer | Linear_complete_layer_RMS_Prop | Linear_complete_layer_momentum | Linear_learner | Linear_learner_bsgd | MDP | Magic_sum | Model_based_reinforcement_Learner | NN | Node | POP_node | POP_search_from_STRIPS | ParticleFiltering | Path | Planning_problem | Plot_env | Plot_prices | Predict | Prob | Q_Learner | RC | RC_DN | RL_agent | RL_env | Rating_set | ReLU_layer | Regression_STRIPS | RejectionSampling | Rob_body | Rob_env | Rob_middle_layer | Rob_top_layer | Runtime_distribution | SARSA_LFA_learner | SLSearcher | STRIPS_domain | SamplingInferenceMethod | Search_from_CSP | Search_problem | Search_problem_from_explicit_graph | Search_with_AC_from_CSP | Searcher | SearcherMPP | Show_Localization | Sigmoid_layer | Simple_game_env | SoftConstraint | State | Strips | Subgoal | TP_agent | TP_env | TabFactor | Updatable_priority_queue | Utility | UtilityTable | VE | VE_DN | Variable | http://aipython.org | Version 0.9.4 | March 2, 2022 |
clause, 87
collaborative filtering, 295
condition, 52
conditional probability distribution (CPD), 177
consistency algorithms, 67
constraint, 52
constraint satisfaction problem, 51
copy_with_assign, 71
CPD (conditional probability distribution), 177
cross validation, 147
CSP, 51
    consistency, 67
domain splitting, 70, 72
search, 65
stochastic local search, 73
currying, 55
data set, 126
DBN
    filtering, 220
unrolling, 220
DBN (dynamic belief network), 216
debbuging, 93
decision network, 225
decision tree learning, 143
decision variable, 225
deep learning, 163
dict_union, 19
display, 17
Displayable, 17
domain splitting, 70, 72
dynamic belief network (DBN), 216
    representation, 216
EM, 253
environment, 21, 22, 267
error, 131
example, 126
explanation, 93
explicit graph, 34
factor, 176, 179
factor_times, 195
feature, 126, 128
feature engineering, 125
file
agentEnv.py, 25
agentMiddle.py, 28
agentTop.py, 29
agents.py, 21
cspConsistency.py, 67
cspDFS.py, 64
cspExamples.py, 55
cspProblem.py, 51
cspSLS.py, 74
cspSearch.py, 65
cspSoft.py, 82
decnNetworks.py, 225
display.py, 17
learnBoosting.py, 157
learnCrossValidation.py, 148
learnDT.py, 143
learnEM.py, 253
learnKMeans.py, 249
learnLinear.py, 150
learnLinearBSGD.py, 156
learnNN.py, 163
learnNoInputs.py, 139
learnProblem.py, 126
logicAssumables.py, 97
logicBottomUp.py, 90
logicExplain.py, 93
logicProblem.py, 87
logicTopDown.py, 92
masLearn.py, 289
masMiniMax.py, 263
masProblem.py, 259
mdpExamples.py, 238
mdpProblem.py, 238
probDBN.py, 217
probDo.py, 222
probFactors.py, 176
probGraphicalModels.py, 180
probHMM.py, 207
probLocalization.py, 211
probRC.py, 188
probStochSim.py, 197
probVE.py, 193
probVariables.py, 175

http://aipython.org

Version 0.9.4 March 2, 2022
<table>
<thead>
<tr>
<th>Function</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>pythonDemo.py</td>
<td>13</td>
</tr>
<tr>
<td>relnCollFilt.py</td>
<td>295</td>
</tr>
<tr>
<td>rlFeatures.py</td>
<td>284</td>
</tr>
<tr>
<td>rlLinExperienceReplay.py</td>
<td>287</td>
</tr>
<tr>
<td>rlModelLearner.py</td>
<td>278</td>
</tr>
<tr>
<td>rlPlot.py</td>
<td>271</td>
</tr>
<tr>
<td>rlProblem.py</td>
<td>267</td>
</tr>
<tr>
<td>rlQExperienceReplay.py</td>
<td>276</td>
</tr>
<tr>
<td>rlQLearner.py</td>
<td>273</td>
</tr>
<tr>
<td>rlQTest.py</td>
<td>275</td>
</tr>
<tr>
<td>rlSimpleEnv.py</td>
<td>269</td>
</tr>
<tr>
<td>rlSimpleGameFeatures.py</td>
<td>281</td>
</tr>
<tr>
<td>searchBranchAndBound.py</td>
<td>47</td>
</tr>
<tr>
<td>searchGeneric.py</td>
<td>41</td>
</tr>
<tr>
<td>searchMPP.py</td>
<td>45</td>
</tr>
<tr>
<td>searchProblem.py</td>
<td>33</td>
</tr>
<tr>
<td>searchTest.py</td>
<td>18</td>
</tr>
<tr>
<td>stripsCSPPlanner.py</td>
<td>114</td>
</tr>
<tr>
<td>stripsForwardPlanner.py</td>
<td>107</td>
</tr>
<tr>
<td>stripsHeuristic.py</td>
<td>109</td>
</tr>
<tr>
<td>stripsPOP.py</td>
<td>118</td>
</tr>
<tr>
<td>stripsProblem.py</td>
<td>101</td>
</tr>
<tr>
<td>stripsRegressionPlanner.py</td>
<td>111</td>
</tr>
<tr>
<td>utilities.py</td>
<td>18</td>
</tr>
</tbody>
</table>

**Index**

- filtering, 209, 213
- DBN, 220
- flip, 19
- forward planning, 106
- frange, 128
- ftype, 128
- game, 259
- Gibbs sampling, 204
- graphical model, 180
- heuristic planning, 109, 113
- hidden Markov model, 207
- hierarchical controller, 25
- HMM
  - exact filtering, 209
  - particle filtering, 213
- HMM (hidden Markov models), 207
- importance sampling, 201
- interact
  - proofs, 94

ipython, 10
k-means, 249
kernel, 136
knowledge base, 88
learner, 138
learning, 125, 161, 163, 173, 249, 257
batched stochastic gradient descent, 156
cross validation, 147
decision tree, 143
deep, 163–173
deep learning, 163
EM, 253
k-means, 249
linear regression, 150
linear classification, 150
neural network, 163
no inputs, 139
reinforcement, 267, 294
relational, 295
supervised, 125, 161
with uncertainty, 249, 257
likelihood weighting, 200
linear regression, 150
linear classification, 150
localization, 211
logistic regression, 178
logit, 151, 152
loss, 131
magic square, 260
magic-sum game, 260
Markov Chain Monte Carlo, 204
Markov decision process, 238
max_display_level, 17
MCMC, 204
MDP, 238, 268
method
  - consistent, 54
  - holds, 53
  - maxh, 109
  - zero, 107
minimax, 259

http://aipython.org

Version 0.9.4
March 2, 2022
minimax algorithm, 263
minsets, 98
model-based reinforcement learner, 278
multiagent system, 259
multiple path pruning, 45
n-queens problem, 62
naughts and crosses, 260
neural network, 163
noisy-or, 179
NotImplementedError, 21
partial-order planner, 117
particle filtering, 201
HMMs, 213
planning, 101, 123, 225–247
CSP, 114
decision network, 225
forward, 106
MDP, 238
partial order, 117
regression, 111
with certainty, 101, 123
with learning, 278
with uncertainty, 225–247
plotting
agents in time, 25
reinforcement learning, 271
robot environment, 30
runtime distribution, 81
stochastic simulation, 205
predictor, 131
Prob, 180
probabilistic inference methods, 187
probability, 175
proof
bottom-up, 90
explanation, 93
top-down, 92
proposition, 87
Python, 9
Q learning, 273
query, 187
queryD0, 222
RC, 188, 232
recursive conditioning, 189
recursive conditioning (RC), 188
recursive conditioning for decision networks, 232
regression planning, 111
reinforcement learning, 267–294
environment, 267
feature-based, 281
model-based, 278
Q-learning, 273
rejection sampling, 199
relational learning, 295
ReLU, 165
resampling, 202
robot
body, 26
environment, 25
middle layer, 28
plotting, 30
top layer, 29
robot delivery domain, 102
runtime, 15
runtime distribution, 81
sampling, 197
importance sampling, 201
belief networks, 198
likelihood weighting, 200
particle filtering, 201
rejection, 199
scope, 52
search, 33
A*, 41
branch-and-bound, 47
multiple path pruning, 45
search_with_any_conflict, 76
search_with_var_pq, 77
show, 54, 182
sigmoid, 151
softmax, 152
stochastic local search, 73
any-conflict, 76
two-stage choice, 77
stochastic simulation, 197

http://aipython.org
tabular factor, 179

test

SLS, 81

tic-tac-toe, 260

top-down proof, 92

uncertainty, 175

unit test, 19, 44, 63, 91, 92, 94

unrolling

DBN, 220

updatable priority queue, 79

utility, 225

utility table, 225

value iteration, 241

variable, 51, 175

variable elimination (VE), 193

variable elimination for decision networks, 236

VE, 193

visualize, 18

XGBoost, 160

yield, 14