Trabalho prático individual nº 2

Inteligência Artificial
Ano Lectivo de 2024/2025

13 de Dezembro de 2024

I Observações importantes

1. This assignment should be submitted via *GitHub* within 28 hours after its publication. The assignment can be submitted after 28 hours, but will be penalized at 5% for each additional hour.

2. Complete the requested methods in module `"tpi2.py"`, provided together with this description. Keep in mind that the language adopted in this course is Python3.

3. Include your name and number and comment or delete non-relevant code (e.g. test cases, print statements); submit only the mentioned module `"tpi2.py"`.

4. You can discuss this assignment with colleagues, but you cannot copy their programs neither in whole nor in part. Limit these discussions to the general understanding of the problem and avoid detailed discussions about implementation.

5. Include a comment with the names and numbers of the colleagues with whom you discussed this assignment. If you turn to other sources, identify these sources as well.

6. All submitted code must be original; although trusting that most students will do this, a plagiarism detection tool will be used. Students involved in plagiarism will have their submissions canceled.

7. The submitted programs will be evaluated taking into account: performance; style; and originality / evidence of independent work. Performance is mainly evaluated concerning correctness and completeness, although efficiency may also be taken into acount. Performance is evaluated through automatic testing. If necessary, the submitted modules will be analyzed by the teacher in order to appropriately credit the student's work.

II Exercícios

Together with this description, you can find modules semantic_network, bayes_net and constraintsearch . They are similar to the ones used in practical lectures, but some changes and additions were introduced.

The attached module `semantic_network` contains new types of relations, namely `AssocSome`, `AssocOne` and `AssoNum`. `AssocSome` is a normal association with no cardinality restrictions. In turn, `AssocOne` allows only one value (e.g. in `hasFather`, each person has only one father).

In module `bayes_net`, mother variables are split into true and false lists.

In module `constraintsearch`, constraint propagation is given implemented.

Module `tpi2` contains some derived classes. In the following exercices, you are asked to complete certain methods/functions in this module. Any other code that you develop and integrate in other modules will be ignored.

The module `tpi2_tests` contains some test code. You can add other test code in this module. Don't change the `semantic_network`, `bayes_net` and `constraintsearch` modules.

You can find the intended results of `tpi2_tests` in the file `tpi2_results.txt`

The responses to the main questions asked by students during this TPI will be collected in section III below.

You are expected to implement the following methods:

1. Implement the method `query(entity,relname)` in class `MySN`, which returns a list of values for a given relation in a given entity, possibly including inherited values. The different types of relations are handled in different ways. In which concerns `Member` and `Subtype` relations, only the local ones are relevant. Inheritance of `AssocOne` and `AssocNum` is processed with cancelling, i.e. the existance of one of these associations in a given entity cancels the effects of a similar association with the same name in a predecessor entity. When there are several declarations of an `AssocOne` association, the most common value should be returned. When there are several declarations of an `AssocNum` association, the average of all values should be returned. In the case of `AssocSome`, inheritance is processed without cancelling. Therefore, all such associations found in predecessors are relevant. Finally, because different users may use a given association name for different types of associations, the type of association most frequently used with a given name should be considered the correct type of the association. For example, `hasHeight` was used four times as `AssocNum` and only once as `AssocOne`, therefore the type of this association is considered to be `AssocNum` and the only declaration in which it was used as `AssocOne` is ignored.

2. Implement the method `test_independence(v1,v2,given)` in class `MyBN`, which determines if two variables, `v1` and `v2`, in a Bayesian network, are conditionally independent of each other when some other variables are given. In order to test independence, the method builds an undirected graph containing the following edges:

   - All edges in the network connecting the input variables (i.e. `v1`, `v2` and those in `given`) to all their ancestors.
   - For each variable covered in the previous point, add edges connecting all pairs of mother variables.
   - As a last step, remove all edges containing any of the variables in `given`.

   The graph is represented as a list of pairs of variables. Variables `v1` and `v2` are independent if there is no path connecting them in the graph. The method receives as input the two variables and a list of other (given) variables, and returns as result a boolean (`True` if independent, false otherwise) and the used graph. The network represented in Figure 1 is included in the `tpi2_tests` module for testing purposes.

3. Implement in class `MyCS` the method `search_all()` which returns all possible solutions without repetitions. This exercice will be evaluated based on correctness, completeness

spark timing (s_t)
$P(s\_t|d) = 0.08$
$P(s\_t|{\sim}d) = 0.002$

distributer (d)
$P(d) = 0.01$

spark plugs (s_p)
$P(s\_p) = 0.3$

car starts (car_s)
$P(car\_s|c\_c \wedge s\_t \wedge s\_q \wedge f\_s) = 0.01$
$P(car\_s|c\_c \wedge s\_t \wedge s\_q \wedge {\sim}f\_s) = 0.081$
$P(car\_s|c\_c \wedge s\_t \wedge {\sim}s\_q \wedge f\_s) = 0.045$
$P(car\_s|c\_c \wedge s\_t \wedge {\sim}s\_q \wedge {\sim}f\_s) = 0.065$
$P(car\_s|c\_c \wedge {\sim}s\_t \wedge s\_q \wedge f\_s) = 0.087$
$P(car\_s|c\_c \wedge {\sim}s\_t \wedge s\_q \wedge {\sim}f\_s) = 0.043$
$P(car\_s|c\_c \wedge {\sim}s\_t \wedge {\sim}s\_q \wedge f\_s) = 0.035$
$P(car\_s|c\_c \wedge {\sim}s\_t \wedge {\sim}s\_q \wedge {\sim}f\_s) = 0.067$
$P(car\_s|{\sim}c\_c \wedge s\_t \wedge s\_q \wedge f\_s) = 0.052$
$P(car\_s|{\sim}c\_c \wedge s\_t \wedge s\_q \wedge {\sim}f\_s) = 0.054$
$P(car\_s|{\sim}c\_c \wedge s\_t \wedge {\sim}s\_q \wedge f\_s) = 0.056$
$P(car\_s|{\sim}c\_c \wedge s\_t \wedge {\sim}s\_q \wedge {\sim}f\_s) = 0.078$
$P(car\_s|{\sim}c\_c \wedge {\sim}s\_t \wedge s\_q \wedge f\_s) = 0.045$
$P(car\_s|{\sim}c\_c \wedge {\sim}s\_t \wedge s\_q \wedge {\sim}f\_s) = 0.031$
$P(car\_s|{\sim}c\_c \wedge {\sim}s\_t \wedge {\sim}s\_q \wedge f\_s) = 0.034$
$P(car\_s|{\sim}c\_c \wedge {\sim}s\_t \wedge {\sim}s\_q \wedge {\sim}f\_s) = 0.023$

voltage at plug (v_p)
$P(v\_p|m\_f \wedge d \wedge b\_v) = 0.003$
$P(v\_p|m\_f \wedge d \wedge {\sim}b\_v) = 0.12$
$P(v\_p|m\_f \wedge {\sim}d \wedge b\_v) = 0.08$
$P(v\_p|m\_f \wedge {\sim}d \wedge {\sim}b\_v) = 0.01$
$P(v\_p|{\sim}m\_f \wedge d \wedge b\_v) = 0.04$
$P(v\_p|{\sim}m\_f \wedge d \wedge {\sim}b\_v) = 0.07$
$P(v\_p|{\sim}m\_f \wedge {\sim}d \wedge b\_v) = 0.13$
$P(v\_p|{\sim}m\_f \wedge {\sim}d \wedge {\sim}b\_v) = 0.09$

spark quality (s_q)
$P(s\_q|s\_p \wedge v\_p) = 0.008$
$P(s\_q|s\_p \wedge {\sim}v\_p) = 0.4$
$P(s\_q|{\sim}s\_p \wedge v\_p) = 0.51$
$P(s\_q|{\sim}s\_p \wedge {\sim}v\_p) = 0.13$

main fuse (m_f)
$P(m\_f) = 0.01$

fuel system (f_s)
$P(f\_s) = 0.1$

battery age (b_a)
$P(b\_a) = 0.5$

battery voltage (b_v)
$P(b\_v|c\_s \wedge b\_a) = 0.18$
$P(b\_v|c\_s \wedge {\sim}b\_a) = 0.02$
$P(b\_v|{\sim}c\_s \wedge b\_a) = 0.90$
$P(b\_v|{\sim}c\_s \wedge {\sim}b\_a) = 0.68$

headlights (h)
$P(h|b\_v) = 0.44$
$P(h|{\sim}b\_v) = 0.89$

alternator (a)
$P(a) = 0.003$

charging system (c_s)
$P(c\_s|a) = 0.48$
$P(c\_s|{\sim}a) = 0.08$

starter system (s_s)
$P(s\_s|s\_m \wedge m\_f \wedge b\_v) = 0.30$
$P(s\_s|s\_m \wedge m\_f \wedge {\sim}b\_v) = 0.21$
$P(s\_s|s\_m \wedge {\sim}m\_f \wedge b\_v) = 0.34$
$P(s\_s|s\_m \wedge {\sim}m\_f \wedge {\sim}b\_v) = 0.12$
$P(s\_s|{\sim}s\_m \wedge m\_f \wedge b\_v) = 0.15$
$P(s\_s|{\sim}s\_m \wedge m\_f \wedge {\sim}b\_v) = 0.13$
$P(s\_s|{\sim}s\_m \wedge {\sim}m\_f \wedge b\_v) = 0.132$
$P(s\_s|{\sim}s\_m \wedge {\sim}m\_f \wedge {\sim}b\_v) = 0.44$

car cranks (c_c)
$P(c\_c|s\_s) = 0.49$
$P(c\_c|{\sim}s\_s) = 0.023$
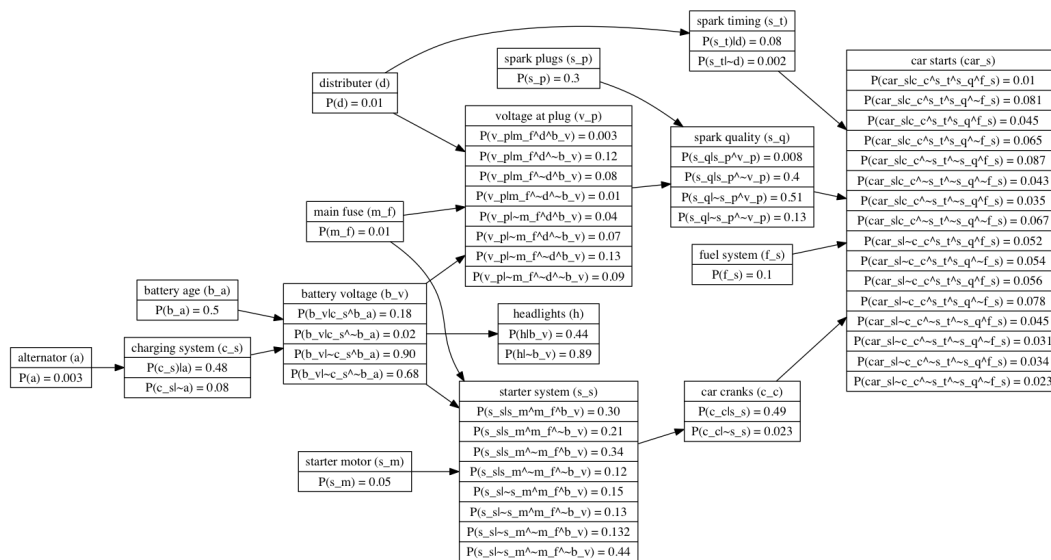
starter motor (s_m)
$P(s\_m) = 0.05$

Figura 1: Bayesian network for a car that does not start

and time efficiency. In order to improve efficiency, consider including some form of variable ordering.

4. Develop the function handle_ho_constraint (domains, constraints , variables , constraint) which, given the **domains** of variables in a problem, the binary **constraints** between these variables, the **variables** involved in a higher-order constraint and a unary **constraint** (which represents the higher-order constraint as presented in classes), will add appropriate auxiliary variable(s) and binary constraint(s) in the constraint graph.

III   Clarification of doubts

This work will be followed through http://detiuaveiro.slack.com. The clarification of the main doubts will be placed here.

1. Podemos adicionar argumentos às funções?

   Resposta: Only optional arguments located after the argments already included; for testing, I will call the functions with the mandatory arguments that are identified in the provided code, and no others.

2. No exercício 2 alguns resultados dão com ordem trocada. É normal? por exemplo: ('c_s', 'b_a') e nos results está ('b_a', 'c_s')

   Resposta: a ordem realmente não é relevante: nem a ordem das arestas na lista, nem a ordem das variáveis em cada aresta; cada aresta (v1,v2) nesse grafo é bi-direccional; basta por num dos sentidos, e vale pelos dois.

3. A ordem nas listas no exercício 1 é importante ou basta aparecerem os valores iguais. Por exemplo, se para **socrates likes** aparecer: ['meat', 'phaenarete', 'milk', 'sophroniscus',

'mathematics', 'philosophy'] em vez de ['philosophy', 'meat', 'sophroniscus', 'phaenarete', 'milk', 'mathematics'].

Resposta: A ordem não é relevante.

4. Is `subtype` meant to return all the subtypes that an entity (of a superclass) has? Or is it the other way round?

   Resposta: Regarding `member` and `subtype`, your code should work as in ex. II.11 of the practical guide; the difference is that you return types (`entity2` in the relation), not full declarations. It does not return subtypes; when you query about `socrates`, it returns an empty list because `socrates` is an instance, not a type, so it cannot be subtype of another type; but it is member of two types; you should look at 'tpi2_results.txt', to better understand what is required.

5. A cotação para cada questão será mais ou menos igual para os quatro exercícios?

   Resposta: As cotações vão ser estas: ex.1 = 30%, ex. 2 = 40%, ex. 3 = 15%, ex. 4 = 15%.

6. Eu não percebo o que são as mães no segundo exercicio E também o grafo temos de ser nós a criar ou ele já existe?

   Resposta: Quando tem uma probabilidade $p(a|r,t)$, r e t são as variáveis mães; no ex. 2, tem que criar um grafo (na forma de uma lista de pares que representam arestas) para decidir sobre a independência das variáveis dadas; veja o que aparece no 'tpi2_results.txt'.

7. No ex. 2, tendo $A \rightarrow B \rightarrow C$ a mãe do C é só o B, ou é o B e o A? E vou buscar as relações ao grafo `self.dependecies`?

   Resposta: A mãe do C nesse exemplo é o B apenas; toda a informação vem originalmente do `self.dependencies`.

8. Are we allowed to import extensions like **from** collections **import** Counter?

   Resposta: You can use anything that is native Python (documented in docs.python.org).

9. Quando temos declarações `AssocNum` é suposto retornar a média dos valores, e se os valores não forem números?

   Resposta: Há realmente duas ocorrências de `AssocNum` em que o valor (segunda entidade) não é um número, mas em que as associações também não são verdadeiras `AssocNum`, mas sim erros do utilizador que declarou; a situação que está a imaginar não ocorre, mas pelo sim, pelo não, se for `AssocNum` e não tiver um número, simplesmente ignoram.

10. Foram corrigidas duas micro-gralhas: faltava a palavra *no* no enunciado do ex. 2 e o `import` do `itertools.product()` no tpi2_tests.py.

11. No exercício 2 os 'ancestors' são a mesma coisa que as 'mothers' ou também incluem as 'mothers' das 'mothers'? E se incuir as 'mothers' das 'mothers', estas devem ser ligadas diretamente às variáveis?

   Resposta: Inclui todas as variáveis ascendentes até chegar às variaveis sem mães; sugiro que veja os resultados dos testes; pode fazer um desenho, para perceber melhor. Cada variavel liga às mães apenas.

12. No exercício 1, como procedemos caso cada tipo de associação apareça a mesma quantidade de vezes, isto é, não houver uma mais commum?

   Resposta: Podem desempatar tendo em conta a seguinte ordem: `AssocSome`, `AssocOne`, `AssocNum`. No entanto, nos casos de teste esse problema nao ocorre.

13. Do enunciado do ex. 2: "For each variable covered in the previous point, add edges connecting all pairs of mother variables." Então vou ligar mães a ancestrais de outras mães?

   Resposta: Para qualquer variável, quer as de entrada, quer as ascendentes, para qualquer uma destas, vai ligar as mães dessa variável entre si: por exemplo, se o primeiro passo identificar as seguintes arestas: $A \rightarrow E, B \rightarrow E, C \rightarrow E, B \rightarrow F, C \rightarrow F, D \rightarrow F, X \rightarrow A, Y \rightarrow A$, vamos ter as seguintes ligações entre mães: $(A, B), (A, C), (B, C), (B, D), (C, D), (X, Y)$, já sem repetições.

14. Podemos definir funções auxiliares?

   Resposta: Podem.

15. Se tivermos: AssocNum('man','hasHeight',1.75)), AssocNum('man','hasHeight',1.80)), AssocNum('man','hasHeight',1.80)) hasHeight deve resultar 1.775 (média dos 2 valores diferentes) ou 1.783 (média dos 3) ?

   Resposta: A média dos três.

16. E se tivermos várias `AssocSome` com `entity2` igual (para a mesma `relname` e `entity1`), é suposto aparecerem todas?

   Resposta: Nesse caso, tiram as repetições.