



# De l'utilisation du dialogue naturel pour masquer les QCM au sein des jeux sérieux



## Résumé

Une des principales faiblesses des jeux sérieux à l'heure actuelle est qu'ils incorporent très souvent des questionnaires à choix multiple (QCM).

Un exemple de QCM :

« Quelles sont les limites des algorithmes évolutionnistes ? »

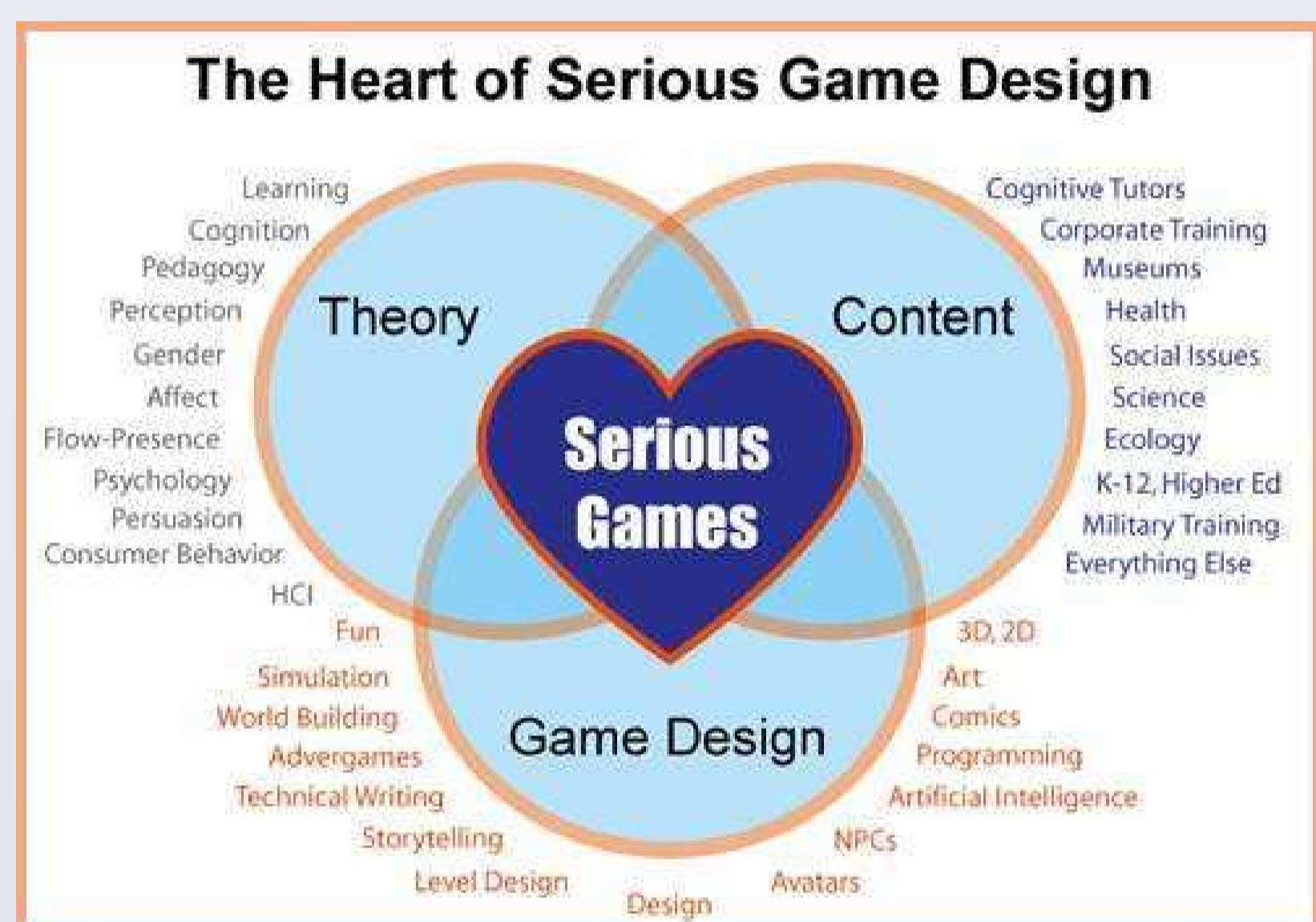
- Ils ne trouveront jamais l'optimal global ;
- Il est en général impossible de faire de l'apprentissage en temps réel avec ;
- Ils ne s'appliquent pas à la plupart des problèmes d'optimisation ;
- Ils n'ont été que peu étudiés dans la littérature.

L'objectif (réponse libre dans le cadre d'un dialogue en langage naturel) :

- Bot : « J'ai entendu dire que les algorithmes évolutionnistes permettraient de résoudre notre problème. Quelles sont les limites de ce genre d'algorithmes ? »
- Humain : « Ces algorithmes prennent beaucoup de temps à s'exécuter et peuvent prendre beaucoup de place en mémoire. »

## État de l'art

Jeux sérieux



Agents conversationnels

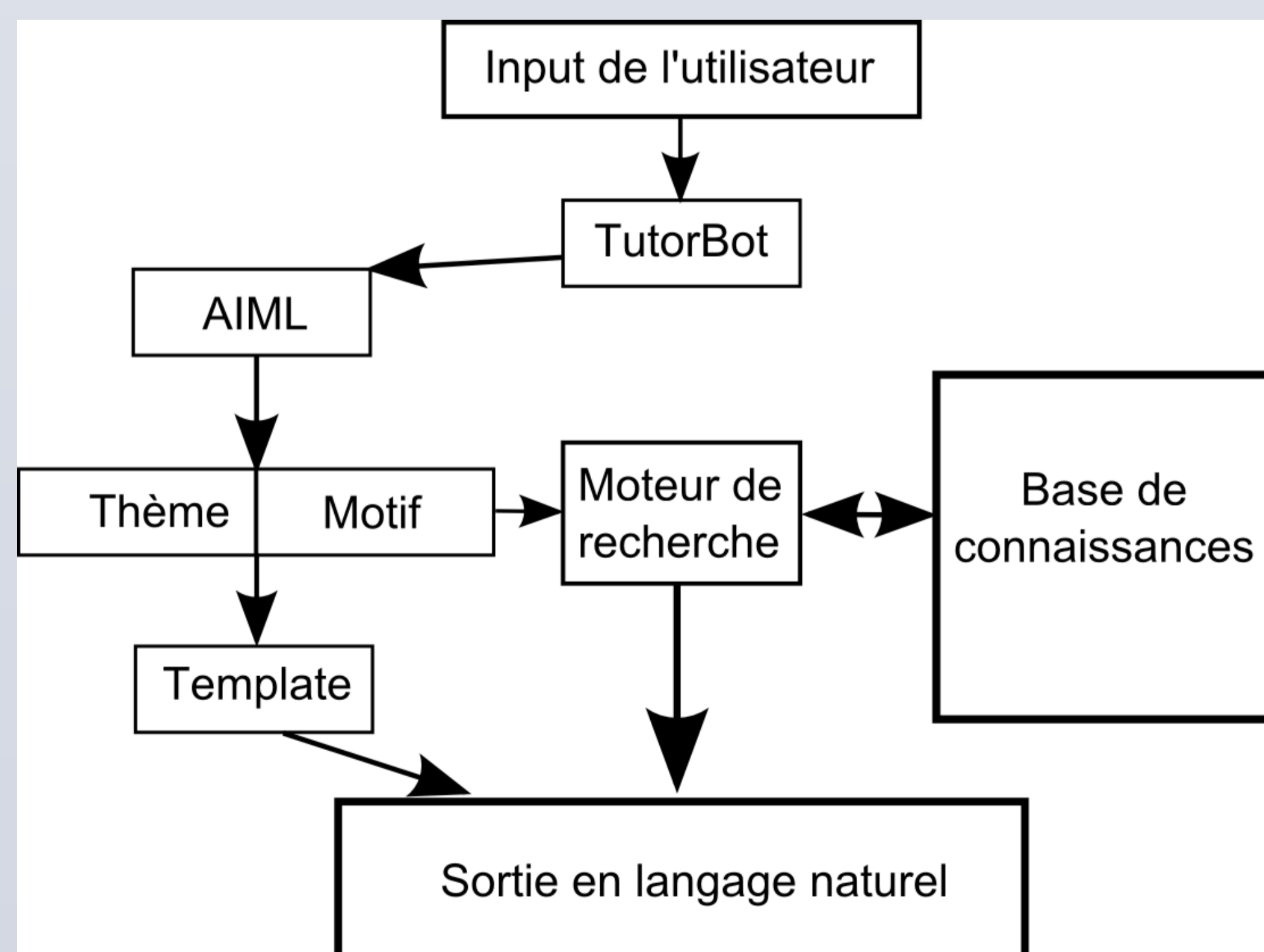
Les agents conversationnels se divisent en deux classes principales :

- les agents conversationnels non orientés tâche,
- les agents conversationnels orientés tâche.

Les agents conversationnels orientés tâche sont eux-mêmes classés usuellement en deux catégories :

- les agents conversationnels orientés service,
- les agents conversationnels éducatifs.

Architecture classique d'un agent conversationnel :



Les limites de l'AIML :

```
<category><pattern>Do I like cheese</pattern>
<template>How would I know if you like cheese or not?</template>
</category>
```

Exemple 1 :  
Human: Hi.  
Cleverbot: How clever are you?  
Human: Very clever.  
Cleverbot: How clever?

Exemple 2 :  
Human: Hi.  
Cleverbot: Are you a boy?  
Human: Yes I am a boys.  
Cleverbot: Are you a girl?

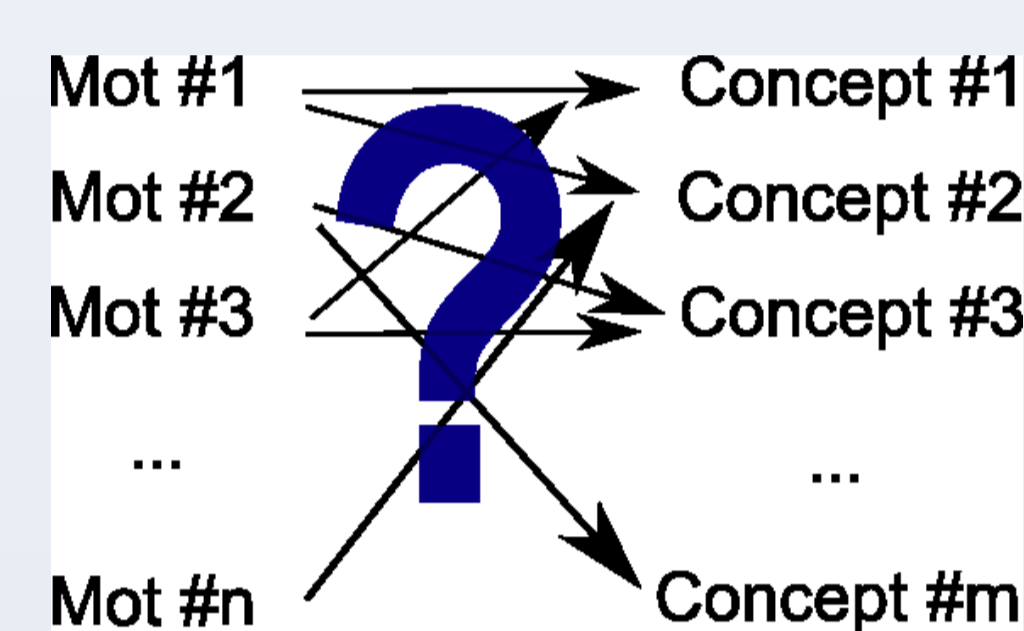
## Les systèmes d'évaluation de réponses libres

Motivés par les résultats expérimentaux montrant :

- les limites des QCM en tant qu'outil d'évaluation de la connaissance des apprenants (Whittington et Hunt, 1999)
- la complémentarité des QCM avec les réponses libres (Anbar, 1991)

1) Project Essay Grade, dirigé par Ellis Batten Page à l'université Duke : utiliser les caractéristiques stylistiques de la réponse tels la taille des mots et le nombre de prépositions (Page, 1968).

2) LSA : comprendre à quel concept le mot fait référence dans le contexte donné.



3) La méthode BLEU : comparer le texte candidat à un ensemble de textes modèles. La note donnée par BLEU au texte candidat se base sur le nombre de N-grammes communs entre le texte candidat et les textes modèles, ce qui s'avère être une mesure efficace malgré sa simplicité, mais est toutefois très sensible au choix d'écriture dans les textes modèles.

## Approche

Particularités de l'approche :

- La réponse de l'apprenant n'est pas notée par rapport à des réponses modèles, mais est reliée à un QCM sous-jacent ;
- Une interaction est possible avec l'apprenant, car le système a la forme d'un agent conversationnel.

Notre système se base sur :

- Analyse syntaxique
- Analyse sémantique (liens avec l'ontologie correspondant au domaine restreint)
- Évaluation (en combinant les diverses techniques de l'état de l'art)
- Apprentissage du système (en enregistrant les réponses des apprenants)

Learning Adventure :



## Conclusions et perspectives

Difficulté :

Une des difficultés dans la recherche de systèmes d'évaluation de réponses libres est l'absence de benchmarks, absence que certains expliquent par des raisons de propriété intellectuelle (Sukkarieh et Blackmore, 2009).

Enjeux :

By year-end 2013, at least 15 percent of Fortune 1000 companies will use a virtual assistant to serve up Web self-service content to enhance their CRM offerings and service delivery. (Gartner 2011)

Éducation en ligne : MITx, Coursera et Udacity.

- cours en ligne gratuits, qui ont déjà attiré plus de 200 000 étudiants ;
- reposent en grande partie sur des QCM pour évaluer les apprenants, à défaut de systèmes plus efficaces.