Project 1
CSE 4/563, Knowledge Representation
Using SNARK as a Propositional Knowledge-Based System
Due 11:00 AM, Wednesday, February 25, 2004

Professor Shapiro

February 4, 2004

1 Introduction

SNARK (SRI's New Automated Reasoning Kit) may be used as a knowledge-based system by performing the following steps.

1. Run Common Lisp.
2. Load `/projects/shapiro/CSE563/snark`
3. Load `/projects/shapiro/CSE563/ask`
4. Change the Common Lisp listener's package to `snark-user`.
5. Load a knowledge base.
6. Ask the system questions.

The first Lisp form in the KB must be `(initialize)`. This must be followed by assertions in the format `(assert` *`wfp`*`)`, where *`wfp`* is a quoted form in the syntax required by SNARK. (*See* [1], Chapter 3.) For an example KB file, see `/projects/shapiro/CSE563/CarPoolWorld.cl`.

An example run of a propositional CarPoolWorld knowledge-based system using SNARK follows. The output following the load of `snark` has been deleted, blank lines and page breaks have been added for readability, and some comments, starting with "*;;;*" have been added.

```
cl-user(1): :ld /projects/shapiro/CSE563/snark
...
; Running on #x80fdafc5 at 2004-02-02T16:26:01

To start using SNARK, change the package to snark-user.
Then use (initialize), (assert <wff>), and (prove <wff>).

cl-user(2): :ld /projects/shapiro/CSE563/ask
; Fast loading /projects/shapiro/CSE563/ask.fasl

cl-user(3): :pa snark-user

snark-user(4): :ld /projects/shapiro/CSE563/CarPoolWorld.cl
; Loading /projects/shapiro/CSE563/CarPoolWorld.cl
; Running on #x80fdafc5 at 2004-02-02T16:26:49
```

```
;;; Ask if ‘‘Tom is the driver or Betty is the driver’’
;;;     is implied by the KB.  (It is.)
snark-user(5): (ask ’(or TomIsTheDriver BettyIsTheDriver))
True

;;; Ask if ‘‘Betty is the driver
;;;          implies that Betty is not the passenger’’
;;;     is implied.  (It is.)
snark-user(6): (ask ’(implies BettyIsTheDriver (not BettyIsThePassenger)))
True

;;; Ask if ‘‘Tom drives Betty and Tom is the passenger’’
;;;     is implied.  (It’s false according to the domain rules.)
snark-user(7): (ask ’(and TomDrivesBetty TomIsThePassenger))
False

;;; Ask if ‘‘Betty drives Tom’’ is implied by the KB.
;;; (Neither it nor its negation is implied.)
snark-user(8): (ask ’(BettyDrivesTom))
Unknown
```

Notice that `(ask` *`wff`*`)` can return `True`, `False`, or `Unknown`. It returns `True` if *`wff`* can be proved from the KB, `False` if `(not` *`wff`*`)` can be proved from the KB, and `Unknown` if there are models that satisfy KB $\wedge$ *`wff`* as well as models that satisfy KB $\wedge\neg$ *`wff`*. We know that `ask` will terminate and return one of these answers because resolution refutation is a decision procedure for propositional logic.

2 The Project

2.1 Project Statement

You are to implement and demonstrate a propositional knowledge-based system about biological organisms using SNARK.

2.2 Background

Biological organisms are categorized in a taxonomy (in computer science, also called a hierarchy or tree). In the “5 kingdoms” model, every living organism is in one of the kingdoms: Animalia, Fungi, Monera, Plantae, Protista, Below the level of Kingdom are the levels Phylum, Class, Order, Family, Genus, and Species. For example, we humans are in the Kingdom Animalia, the Phylum of Chordates, the Class of Mammals, the Order of Primates, the Family of Homonids, the Genus of Homo, and the Species of sapiens. Protozoa and Algae are in Kingdom Protista; molds and yeasts are in Kingdom Fungi.

An important property of hierarchies is that the subcategory relation is transitive. That is, since Homo sapiens are a subcategory of Hominids, and Homonids are a subcategory of Primates, then Homo sapiens are a subcategory of Primates.

In set theory, if sets b, c, and d **partition** the set a, then

1. anything that is in b, c, or d is in a (b, c, and d are subsets of a);
2. if anything is in a, then it’s in b, c, or d (b, c, and d exhaust a);
3. if anything is in b, c, or d, it’s not in either of the others (b, c, and d are mutually exclusive).

For example, the 5 kingdoms partition the set of living organisms.

Besides the 5 kingdom model, animals may be partitioned into herbivores, carnivores, and omnivores. Humans, for example are omnivores (even though some humans are vegetarians).

Animals may also be partitioned according to their principle mode of transportion: on land, through the water, or through the air. For example, dolphins are a Family of the Order of Cetaceans of the Class of Mammals, whose principal mode of transportation is through the water.

2.3 The Task

You are to create a knowledge base of living organisms that includes at least the 21 categories: Animalia, Fungi, Monera, Plantae, Protista, Chordates, Mammals, Primates, Homonids, Homo, sapiens, protozoa, algae, molds, yeasts, herbivores, carnivores, omnivores, moves on land, moves through the water, moves through the air, and at least 30 others, which you can find from any reference source. Besides the partitions of the 5 kingdoms, the 3 diets, and the 3 modes of transportation, your KB must include at least 3 other partitions. Other subcategories may be non-exhaustive, but your KB must contain the information that two categories are exclusive whenever they really are.

You are to include a demonstration of your system that shows the `ask` function answering questions that you devise. The questions you demonstrate should be carefully chosen to show off your KB and the abilities of `ask` and SNARK. They must include at least one question answered `True`, at least one answered `False`, and at least one answered `Unknown`. If A, B, and C are wfps, not necessarily atomic, your questions must include at least one of each of the following forms for which the answer is `True`: $A \Rightarrow B$, $A \Rightarrow \neg B$, $A \wedge \neg B \Rightarrow C$, $A \Rightarrow B \vee C$. They must include questions that show that your KB is designed correctly for subsets, exhaustive sets, and mutually exclusive sets, and that you can make proper use of combinations of taxonomic information, dietary information, and mode-of-transportation information. As much as possible, your questions should require your system to use the transitivity property of the KB.

2.4 Deliverables

As it says on the CSE 4/563 web page, you are to hand in a conference-style paper, produced using a document formatting program such as Microsoft Word or LaTeX, and printed on 8.5 by 11 inch paper, stapled in the upper left-hand corner, with your own title page , plus a well-documented listing and run of your program. (Do not enclose your paper in a folder or cover.) The paper is due at the start of class, 11:00 AM, Wednesday, February 25, 2004.

In addition to the paper, you are to `submit` (using `submit_cse463` or `submit_cse563`) your program, so that it can be run and checked if the instructors choose. You are to submit your program soon enough before 11:00 AM, Wednesday, February 25, 2004, so that you can get to class by 11:00 to hand in the paper.

Note that "your program" is the KB file and the sequence of calls to `ask`. In fact, for ease of testing and debugging, your KB file can end with the sequence of calls to `ask`. Name your file `proj1KB.`*`ext`*, for some *`ext`*.

2.4.1 The Paper

Your paper should have the following parts:

1. Descriptive title
2. Author identification
3. Introduction: general description of the project
4. Domain: English description of the domain your program works in
5. Formalism
   (a) Syntax and semantics of propositions
   (b) English and formal presentation of information in the KB
6. Demonstration run, properly commented and formatted
7. Acknowledgments as needed
8. References as needed

2.5 Grading

Project grading will be according to the following table.

	CSE 463	CSE 563
KB		
KB has the 21 required categories	6	2
KB has at least 30 other categories	6	3
KB has the 3 required partitions	6	2
KB has at least 3 other partitions	6	4
KB has some exclusive, non-exhaustive categories	5	4
Correct representation of subsets	2	3
Correct representation of exhaustive subsets	2	3
Correct representation of mutually exclusive subsets	2	3
KB Subtotal	**35**	**24**
Questions		
At least one False question	3	2
At least one Unknown question	3	2
True question of form $A \Rightarrow B$	3	3
True question of form $A \Rightarrow \neg B$	3	3
True question of form $A \wedge \neg B \Rightarrow C$	3	3
True question of form $A \Rightarrow B \vee C$	3	3
Questions show subsets	3	4
Questions show exhaustive sets	4	4
Questions show mutually exclusive sets	4	4
Questions show combinations of taxonomy/diet/transportation	3	4
A substantial number of questions requires the use of transitivity	3	4
Questions Subtotal	**35**	**36**
Paper		
Paper format	2	3
General project description	4	6
English description of domain	5	6
Syntax and semantics of propositions	6	7
Description of formalization of KB	6	7
Presentation of demonstration run	5	7
Acknowledgments and References	2	4
Paper Subtotal	**30**	**40**
Project Total	**100**	**100**

Acknowledgments

Prof. Shapiro appreciates the advice of Caren D. Shapiro, PhD, Associate Professor of Biology at D'Youville College, on biological taxonomies, and of Richard Waldinger, SRI International, on the SNARK function that flushes clauses generated after the assertions.

References

1. Mark E. Stickel, Richard J. Waldinger, & Vinay K. Chaudhri, "A Guide to SNARK", `http://www.ai.sri.com/snark/tutorial/tutorial.html`
2. Montgomery College Library, Guide to Internet Resources for Biological Taxonomy & Classification `http://mclibrary.nhmccd.edu/taxonomy/taxonomy.html`
3. Classification/Taxonomy of Cetaceans, `http://www.crru.org.uk/education/factfiles/taxonomy.htm`