

TRƯỜNG ĐẠI HỌC SÀI GÒN
KHOA CÔNG NGHỆ THÔNG TIN



Homework 3
HỌC PHẦN: TRÍ TUỆ NHÂN TẠO NÂNG CAO

Giảng viên hướng dẫn : Ts. Đỗ Như Tài

Nhóm 11:

Phạm Văn Nam

3122410251

Thành phố Hồ Chí Minh - Tháng 11/2025

Mục lục

Câu 1 Discussion-Based Warm Ups.....	3
Câu 2 Wandering in Wumpus World.....	5
Câu 3 Axioms & Arrows	7

Câu 1 Discussion-Based Warm Ups

(a) Vấn đề về con kỳ lân (The Unicorn Problem)

Ký hiệu các mệnh đề:

- M: Con kỳ lân là sinh vật huyền thoại (mythical).
- I: Con kỳ lân là bất tử (immortal).
- A: Con kỳ lân là động vật có vú (mammal).
- H: Con kỳ lân có sừng (horned).
- G: Con kỳ lân có phép thuật (magical).

Các thông tin được cho:

1. Nếu con kỳ lân là sinh vật huyền thoại thì nó bất tử. $M \rightarrow I$
2. Nếu con kỳ lân không phải sinh vật huyền thoại thì nó là động vật có vú và phải chết. $\neg M \rightarrow (\neg I \wedge A)$
3. Nếu con kỳ lân hoặc bất tử hoặc là động vật có vú thì nó có sừng. $(I \vee A) \rightarrow H$
4. Nếu con kỳ lân có sừng thì nó có phép thuật. $H \rightarrow G$

Lập luận suy luận

Xét hai khả năng có thể xảy ra:

1. Trường hợp 1: Con kỳ lân là sinh vật huyền thoại ($M = \text{True}$)
 - Theo (1): Nó bất tử ($I = \text{True}$).
 - Khi đó $(I \vee A) = \text{True} \Rightarrow$ theo (3): Nó có sừng ($H = \text{True}$).
 - Theo (4): Nó có phép thuật ($G = \text{True}$).
2. Trường hợp 2: Con kỳ lân không phải sinh vật huyền thoại ($M = \text{False}$)
 - Theo (2): Nó là động vật có vú ($A = \text{True}$) và phải chết ($I = \text{False}$).
 - Khi đó $(I \vee A) = \text{True} \Rightarrow$ theo (3): Nó có sừng ($H = \text{True}$).
 - Theo (4): Nó có phép thuật ($G = \text{True}$).

Kết luận

Tính chất	Có thể suy ra chắc chắn?	Giải thích
Có sừng (H)	Có thể suy ra chắc chắn	Dù trong trường hợp nào, $(I \vee A) \rightarrow H$ luôn đúng
Có phép thuật (G)	Có thể suy ra chắc chắn	Vì $H \rightarrow G$ và H luôn đúng
Là sinh vật huyền thoại (M)	Không xác định	Cả hai khả năng $M = \text{True}$ hoặc False đều phù hợp với dữ kiện

→ Kết luận:

Con kỳ lân chắc chắn có sừng và có phép thuật, nhưng không thể kết luận chắc chắn rằng nó là sinh vật huyền thoại.

(b) Bài tập về quan hệ hệ quả logic (Entailment)

Phân tích từng mệnh đề xem có đúng về mặt logic hay không:

STT Mệnh đề	Kết quả	Giải thích
(i) $(A \wedge B) \models (A \Leftrightarrow B)$	Sai	Khi A và B cùng đúng thì $(A \wedge B)$ đúng, nhưng điều đó không đảm bảo rằng $A \Leftrightarrow B$ luôn đúng trong mọi trường hợp.
(ii) $A \Leftrightarrow B \models A \vee B$	Đúng	Nếu $A \Leftrightarrow B$ thì hai biến có cùng giá trị; nếu cùng đúng, $A \vee B$ cũng đúng.
(iii) $A \Leftrightarrow B \models \neg A \vee B$	Đúng	Nếu $A \Leftrightarrow B$ thì hai biến có cùng giá trị, nên biểu thức $\neg A \vee B$ luôn đúng trong mọi trường hợp.
(iv) $(A \wedge B) \rightarrow C \models (A \rightarrow C) \vee (B \rightarrow C)$	Đúng	Nếu C đúng khi cả A và B đúng, thì ít nhất một trong hai điều kiện $A \rightarrow C$ hoặc $B \rightarrow C$ phải đúng.
(v) $(A \vee B) \wedge \neg(A \rightarrow B)$ là mệnh đề khả thỏa (satisfiable)	Đúng	Chọn $A = \text{True}$, $B = \text{False} \Rightarrow (A \vee B) = \text{True}$ và $(A \rightarrow B) = \text{False}$, do đó mệnh đề khả thỏa.

(c) Phân biệt giữa “Tính khả thỏa” và “Hệ quả logic”

- Tính khả thỏa (Satisfiability):
Kiểm tra xem có ít nhất một mô hình (thế giới khả dĩ) mà trong đó công thức là đúng hay không.
→ Câu hỏi đặt ra: “Công thức này có thể đúng được không?”
- Hệ quả logic (Entailment – \models):
Kiểm tra xem một công thức luôn đúng trong mọi mô hình mà tập giả thiết (premises) đúng.
→ Câu hỏi đặt ra: “Từ các giả thiết này, công thức kia có bắt buộc phải đúng không?”

Ví dụ minh họa

- Mệnh đề “Trời mưa hoặc trời nắng” là khả thỏa, vì có thể xảy ra trong một thế giới nào đó.
- Nếu “Trời mưa” \models “Mặt đất ướt” thì có nghĩa là trong mọi thế giới mà trời mưa, mặt đất luôn ướt.

Tóm lại:

- Tính khả thỏa (Satisfiability) \rightarrow “có thể đúng”.
- Hệ quả logic (Entailment) \rightarrow “phải đúng trong mọi trường hợp”.

Câu 2 Wandering in Wumpus World

(a) Entailment và Satisfiability trong Wumpus World

Trong Wumpus World, tác tử (agent) phải suy luận về trạng thái của các ô trong bản đồ dựa trên các cảm nhận (percepts) như Gió (Breeze) và Mùi hôi (Stench).

Để mô tả quy tắc logic của thế giới này, ta sử dụng các mệnh đề logic mệnh đề (propositional logic) sau đây:

Các mệnh đề cơ bản của Thế giới Wumpus

1. Liên hệ giữa Gió (Breeze) và Hố (Pit):
$$\text{Breeze}(x, y) \Leftrightarrow \text{Pit}(x \pm 1, y) \vee \text{Pit}(x, y \pm 1)$$

Có gió tại ô (x, y) khi và chỉ khi ít nhất một ô kề nó có hố.

2. Liên hệ giữa Mùi hôi (Stench) và Wumpus:
 $\text{Stench}(x, y) \Leftrightarrow \text{Wumpus}(x \pm 1, y) \vee \text{Wumpus}(x, y \pm 1)$
 \rightarrow Có mùi hôi tại ô (x, y) khi và chỉ khi ít nhất một ô kề có Wumpus.
3. Khi không có gió:
 $\neg \text{Breeze}(x, y) \Rightarrow \neg \text{Pit}(x \pm 1, y) \wedge \neg \text{Pit}(x, y \pm 1)$
 Nếu không có gió tại (x, y) thì các ô kề không có hố.
4. Khi không có mùi hôi:
 $\neg \text{Stench}(x, y) \Rightarrow \neg \text{Wumpus}(x \pm 1, y) \wedge \neg \text{Wumpus}(x, y \pm 1)$
 Nếu không có mùi hôi tại (x, y) thì các ô kề không có Wumpus.
5. Định nghĩa ô an toàn:
 $\text{Safe}(x, y) \Leftrightarrow \neg \text{Pit}(x, y) \wedge \neg \text{Wumpus}(x, y)$
 Một ô được xem là an toàn khi và chỉ khi nó không chứa hố và không chứa Wumpus.

Phân tích theo Figure 3 (Entailment vs Satisfiability)

Giả sử tác tử có các cảm nhận sau:

- Tại ô (1,1): không có gió, không có mùi hôi $\rightarrow \neg \text{Breeze}(1,1) \wedge \neg \text{Stench}(1,1)$
- Tại ô (2,1): có gió và có mùi hôi $\rightarrow \text{Breeze}(2,1) \wedge \text{Stench}(2,1)$

Từ đó, ta áp dụng các mệnh đề logic trên để suy luận (entailment) và xác định các khả năng (satisfiability):

Ô	Entailment (chắc chắn)	Satisfiable (có thể)	Giải thích
(1,1) Safe	-		Không có gió, không có mùi hôi \rightarrow các ô kề đều an toàn
(1,2) Safe	-		Kề (1,1) không có gió \rightarrow không có hố ở (1,2)
(2,1) -		Có thể có Pit hoặc Wumpus	Breeze và Stench \rightarrow có thể tồn tại cả Pit và Wumpus gần đó
(2,2) -		Có thể có Pit	Breeze ở (2,1) cho phép Pit ở (2,2)

Ô	Entailment (chắc chắn)	Satisfiable (có thể)	Giải thích
(3,1) Wumpus	-		Có Stench tại (2,1) và không có Stench ở (1,1) \Rightarrow Wumpus nằm tại (3,1)

(b) Liên hệ với mã giả của tác tử (Hybrid-Wumpus-Agent)

Ghép trạng thái trong Figure 4 với mã giả

Trạng thái trong hình	Khối mã tương ứng	Giải thích
Hình 1	Khối D	Vì hết ô an toàn, cần mạo hiểm.
Hình 2	Khối B	Agent chuẩn di khám đến để ô chắc chắn được cho là an toàn (những ô chưa thăm).
Hình 3	Khối A	Phát hiện vàng.
Hình 4	Khối E	Không thể suy luận thêm, chuẩn bị thoát

Câu 3 Axioms & Arrows

(a) Successor-State Axioms là gì?

Action Axioms là ta chỉ mô tả một chiều thay đổi, ví dụ:

Nếu bắn \rightarrow không còn mũi tên.

Nhưng như vậy còn thiếu, vì ta chưa nói:

“Nếu không bắn, thì vẫn còn mũi tên.”

Thế là người ta tạo ra Successor-State Axiom (SSA), để gom cả hai điều đó vào một câu duy nhất, ví dụ:

$$\text{HaveArrow}_{t+1} \Leftrightarrow (\text{HaveArrow}_t \wedge \neg \text{Shoot}_t)$$

Câu này nghĩa là:

Ở bước kế tiếp ($t+1$), bạn có mũi tên nếu và chỉ nếu ở bước hiện tại (t) bạn đang có mũi tên và không bắn. \rightarrow Rất gọn, vì nó bao gồm cả thay đổi và giữ nguyên.

Đây chính là cách giải quyết “Frame Problem” — tức là không cần liệt kê hàng trăm thứ “nếu không đụng tới thì vẫn y nguyên” nữa.

(b) Nếu làm hành động khi điều kiện không thoả mãn thì sao

Ví dụ:

Giả sử ta không có mũi tên, mà vẫn làm hành động Shoot.

Theo công thức SSA:

$$\text{HaveArrow}_{t+1} \Leftrightarrow (\text{HaveArrow}_t \wedge \neg \text{Shoot}_t)$$

Thay vào:

$$\text{HaveArrow}_t = \text{False}, \text{Shoot}_t = \text{True}$$

Về phải là $(\text{False} \wedge \neg \text{True}) = (\text{False} \wedge \text{False}) = \text{False}$.

Vậy:

$$\text{HaveArrow}_{t+1} \Leftrightarrow \text{False}$$

Nghĩa là: Ở bước sau, mình vẫn không có mũi tên.

Tức là:

Nếu làm hành động không hợp lệ (bắn khi không có tên), thì thế giới không thay đổi gì cả.

(c) Nếu kế hoạch có hành động sai thì sao

Một kế hoạch (plan) được xem là “đúng” nếu từ trạng thái ban đầu + các luật SSA, ta có thể suy ra rằng mục tiêu (goal) cuối cùng sẽ đạt được.

Nhưng nếu trong kế hoạch đó có hành động sai (ví dụ như “bắn” khi không có tên, hay “mở cửa” khi chưa có chìa khoá), thì như phần (b) đã chỉ ra — hành động đó không có tác dụng gì.

Kết quả là toàn bộ kế hoạch không thể đạt được mục tiêu, vì một số bước “vô nghĩa”.

