RMIT UNIVERSITY Computer Science and Information Technology

Expectation Reasoning and Agent's Observations

Bình Vũ Trần (tvubinh@cs.rmit.edu.au)
James Harland (jah@cs.rmit.edu.au)
Margaret Hamilton (mh@cs.rmit.edu.au)

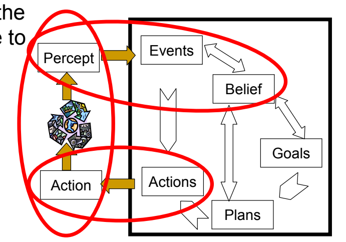
Where should I fly?

- An eagle chasing a sparrow in a cave system usually faces different options where a decision needs to be made. Questions such as
  - Should I follow this path or stay in this cave?
  - I cannot see the sparrow here, should I fly to the next cave or go back?

need to be answered.

BDI Agent vs. Environment

- Representation of the real world structure to assist reasoning
- The ability to make assumption when information is unavailable



(Winikoff, 2001)

Why Observation?

- Observation: "An act of seeing or of fixing the mind upon anything"
  - Not only obtaining information through sensors
  - But also a **process – observation method –** that uses effectors to position sensors or other effectors
- Such process can be nested and hence hierarchical

Why Observation (cont.)?

- Why not "perception"?
  - Only recognition and interpretation of sensory stimuli
  - Do not explicitly consider sensory settings
- Why not "action"?
  - Only about the effects produced
  - Missing the link of how an effect is captured by a sensory stimulus

Observation Reasoning

- World structure is captured using
  - Sensors: which directly bring the images of an environment to the agent's mind
  - Effectors: which **position** the agent and its sensors in the environment
  - Methods: which determine how sensors and effectors should be used together
- Observation reasoning: Reasoning about how to capture the world structure through observations

Why Agent Should Make Assumptions?

- Due to the agent's limited capabilities
  - Cannot always make observations
  - Cannot make observations simultaneously
- Assumptions are used
  - To delay justifications by direct observations
  - To incorporate with others to recover simultaneity view

RMIT UNIVERSITY

Reasoning with Assumptions

- **Transparency**: Assumptions should utilise the same processes as observation reasoning
- **Saliency**: Assumption should elicit the same behaviour response as equivalent physical concept

RMIT UNIVERSITY

Why Expectation?

- Expectation
  - "is a mental act and has always a reference to … some coming event."
  - A mental state associated directly with an observation. Hence, experience-oriented.
  - Unavailable observation can be postponed and verified later. Expectation enables an agent to make assumption.

RMIT UNIVERSITY

Why Expectation (cont.)

- Why not knowledge? $K\psi \Rightarrow \psi$
  - Locally resource-bounded cannot be perfectly synchronised with the environment.
  - Changes in the environment can be much faster or slower than the agent's capabilities of capturing them.
- Why not belief?
  - Justification problem to classify true/false beliefs

RMIT UNIVERSITY

Expectation Reasoning

- Concerns about
  - The results of observations
  - The process of observations and substitutions of assumptions by ground observation
- Reason about expectation violations
  - Grounding error: once assumptions cannot be justified using the given substitutions
  - **Difference in values**: true and false

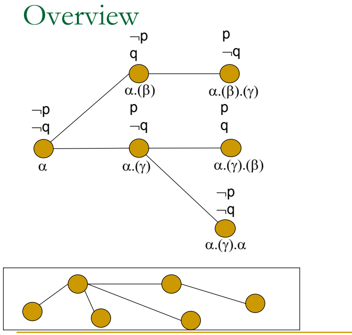
RMIT UNIVERSITY

Agents in Observation System

- An agent's set of sensors and effectors is fixed
- Observation can be extended by incorporating sensors and effectors but must be rooted to the set of sensors
- By taking further observations, the agent associates its expectation with the observations
- Hypothetical effect is justified by a sensory information
- Every observation in a sequence must be well-defined to be a strong observation sequence.

RMIT UNIVERSITY

Observation-Expectation System – An Overview



- α: look
- β: flap
- γ: turn head

- p: sparrow
- q: dark

Observation Methods

- Linear observation method: is a sequence σ of observations where
  - ς, the observation from innate set of sensors is an observation sequence
  - If σ is a sequence then so are σ.η and σ.(η) where η from sensors, (η) from effectors
- Given a free set of assumptions $\Psi = \{x, y, \ldots\}$
  - If σ is a sequence then so is σ.(*x*) but not σ.*x*

Observation Interpretation – Possible Worlds Mapping

- Observation interpretation function
  I: $\Gamma \rightarrow G_i \cup \{\bot\}$
- Conditions
  - Justification: I(σ.(η)) = I(σ.η)
  - Individuality: $I(\eta) \in G_i$, $\forall \eta \in S_i$
  - Expectability: τ = σ.[η], $I(\sigma) \in G_i$, $I(\tau) \in G_i$ iff $I(\sigma) \sim_e I(\tau)$
  - Entireness: if I(τ) = ⊥ for some τ = prefix(σ) then I(σ) = ⊥

Semantics – Expectation Model

- Observation Naming: N: $\Xi \rightarrow \Gamma$
- Expectation language:
  $wff ::= s \mid p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \Rightarrow \psi \mid \langle E_i \rangle \varphi \mid [E_i]\varphi \mid @_s \varphi$
- Kripke model M = $\langle G_i, \sim_e, \pi \rangle$
  - Expectation function $\pi: \Xi \cup \Phi \rightarrow \wp(G_i)$
  - $\langle M, g \rangle \models s$ iff $\pi(s) = \{g\}$ for all $s \in \Xi$
  - $\langle M, g \rangle \models @_s\varphi$ iff $\langle M, g_s \rangle \models \varphi$ where g_s is a denotation of s

Inference Rules – **KE** System

$\dfrac{\begin{array}{c}@_s\varphi \vee \psi \\ \neg @_s \varphi\end{array}}{@_s\psi}$ $\quad$ $\dfrac{\neg @_s\varphi \vee \psi}{\begin{array}{c}\neg @_s \varphi \\ \neg @_s\psi\end{array}}$ $\quad$ $\dfrac{\begin{array}{c}\neg @_s\varphi \wedge \psi \\ @_s \varphi\end{array}}{@_s\psi}$ $\quad$ $\dfrac{@_s\varphi \wedge \psi}{\begin{array}{c}@_s \varphi \\ @_s\psi\end{array}}$

$\dfrac{\begin{array}{c}@_s\varphi \Rightarrow \psi \\ @_s \varphi\end{array}}{@_s\psi}$ $\quad$ $\dfrac{\begin{array}{c}@_s\varphi \Rightarrow \psi \\ \neg @_s \psi\end{array}}{\neg @_s\varphi}$ $\quad$ $\dfrac{\neg @_s\varphi \Rightarrow \psi}{\begin{array}{c}@_s \varphi \\ \neg @_s\psi\end{array}}$ $\quad$ $\dfrac{@_s \neg\psi}{\neg @_s \psi}$ $\quad$ $\dfrac{\neg @_s \neg\psi}{@_s \psi}$

$\dfrac{@_s @_t \psi}{@_t \psi}$ $\quad$ $\dfrac{\neg @_s @_t \psi}{@_t \psi}$ $\quad$ $\dfrac{@_s t}{@_t s}$ $\quad$ $\dfrac{@_s t \quad @_s \psi}{@_t \psi}$ $\quad$ $\dfrac{@_s \langle E \rangle t \quad @_t \psi}{@_s \langle E \rangle \psi}$

$\dfrac{\neg @_s \langle E \rangle \psi \quad @_s \langle E \rangle t}{\neg @_t \psi}$ $\quad$ $\dfrac{@_s [E] \psi \quad @_s \langle E \rangle t}{@_t \psi}$ $\quad$ $\dfrac{@_s \langle E \rangle \varphi}{\begin{array}{c}@_s \langle E \rangle a \\ @_a \psi\end{array}}$ $\quad$ $\dfrac{\neg @_s [E] \varphi}{\begin{array}{c}@_s \langle E \rangle a \\ \neg @_a \psi\end{array}}$

(D'Agostino 1994)

Expectation Violations and Inference Rules

$\overline{@_t \psi \mid \neg @_t \psi}$

- Difference in values
  - Based on the Principle of Bivalence: every proposition is either true or not true.
  - Hence, if the original expectation is satisfiable then any unexpected observation should not invalidate the expectation.
- Grounding error

Observation Model Construction Steps

- Finding equivalent class of observation sequences
- Building up accessibility relation between equivalent classes by adding or removing observations
- Valuation function returns all equivalent classes that a proposition can be observed.

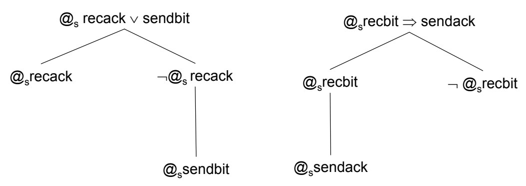
Case study: Bit transmission problem

- System settings
  - A sender A
  - A receiver B
  - A communication medium C
- It is trivial if A, B, and C are all reliable.
- Any fault with any combination of A, B, C requires a new protocol to guarantee delivery

Case 1: Communication Line Unreliability

- Communication line is faulty which can drop message (but not flip the bit)
- Simple protocol:
  - A sends its bit until receives an acknowledgment
  - B sends acknowledgment when receives a bit

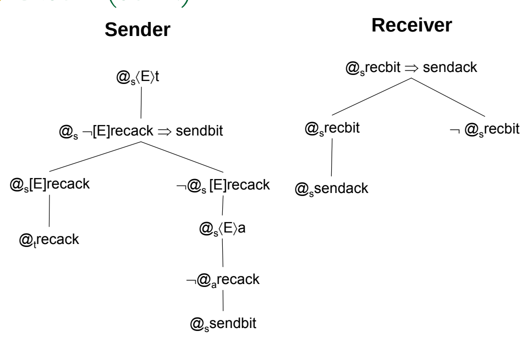
Case 1 (cont.)



Case 2: Communication Line Unreliability

- Communication line: delivery time is delayed after a **fixed period *k***, a message can be dropped
- A protocol can be defined with some additional information to construct the protocol
  - $|t| = |s| + k$
  - $|a| = |s| + k + 1$

Case 2 (cont.)



Coordinated Attack Problem (Gray, 1978)

- Two divisions of an army camped on two hilltops overlooking a valley with the enemy awaiting.
- Both divisions must attack the enemy simultaneously to win the battle.
- The only means of communication is by messengers
- It takes 1 hour to deliver a message.

Impossibility of Coordinated Attack

- Yemini & Cohen (1979), Halpern & Moses (1990) shows impossibility of the problem for the strong requirement: "The coordinated attack must be **simultaneous"**
- Dilemma:
  - Common knowledge must involve simultaneous change to achieve coordination
  - Simultaneity cannot be attained in practice

Towards Agent Coordination

- Common Knowledge
  - Eventual Common Knowledge
  - Time-stamped Common Knowledge
- Common Expectation
  - Ability to delay justification
  - Naming mechanism allows representation like timestamps

Towards a Flexible Protocol Specification

- An expectation violation occurs, the principle of bivalence is applied. Flexibility of the model construction depends on this.
- A fault-tolerant protocol is the one that is independent of PB application.

Other Concepts Related to Expectation

- Commitment
- Obligation
- Reputation/Trust
- Confidence
- Intention

System Characteristics

- Ground expectation reasoning system is decidable.
- Observation layer allows representation of observations in agent's control and other observations from the environment.
- An infrastructure for a scalable system.

Further work

- Expectation Interpreter/Reasoner Implementation
- Observation method construction
- Combining observation methods in MAS using fibring analytic tableaux
- Agent cooperation based on common expectation

Thank you

???