

# Introduction to Artificial Intelligence



10S3001 - Artificial Intelligence

Samuel I. G. Situmeang

**Faculty of Informatics and Electrical Engineering**

# Objectives

Students are able:

- to explain the definition of artificial intelligence,
- to describe various examples of artificial intelligence technology in everyday life,
- to describe examples of intelligent agents, and
- to describe the types and environments of intelligent agents.



1053001-AI | Institut Teknologi Del

2

Siswa mampu:

- menjelaskan pengertian kecerdasan buatan,
- mendeskripsikan berbagai contoh teknologi kecerdasan buatan dalam kehidupan sehari-hari,
- menjelaskan contoh agen cerdas, dan
- menggambarkan jenis dan lingkungan agen cerdas.

# Reasons to Study AI

- To learn more about ourselves.
- The constructed intelligent entities are interesting and useful in their own right.
- AI still has openings for one to contribute new ideas.
- AI currently encompasses a huge variety of subfields.

1053001-AI | Institut Teknologi Del



Kecerdasan manusia adalah fenomena yang kompleks dan masih belum sepenuhnya dipahami. Dengan mempelajari kecerdasan buatan, kita dapat mengembangkan pemahaman yang lebih baik tentang cara kerja kecerdasan manusia. Semakin kita memahami cara kerja kecerdasan manusia, maka teknologi yang dikembangkan dalam berbagai bidang kehidupan akan semakin baik.

Di samping itu, kecerdasan buatan memiliki potensi untuk mengubah cara kita hidup dan bekerja. Entitas cerdas yang dibangun dapat membantu kita menyelesaikan tugas-tugas kompleks, membuat keputusan yang lebih baik, dan memahami dunia di sekitar kita dengan lebih baik. Beberapa contoh bagaimana entitas cerdas yang dibangun dapat menarik dan berguna adalah:

- Robotik : Robot dapat digunakan untuk melakukan tugas-tugas yang berbahaya atau membosankan bagi manusia, seperti menambang, manufaktur, dan perawatan kesehatan.
- Asisten virtual : Asisten virtual seperti Siri dan Alexa menggunakan kecerdasan buatan untuk membantu pengguna dengan berbagai tugas, seperti mengatur jadwal, membuat catatan, dan menjawab pertanyaan.

Bidang kecerdasan buatan masih berkembang pesat, dan ada banyak peluang untuk berkontribusi pada penelitian dan pengembangan. Beberapa contoh bagaimana seseorang dapat menyumbangkan ide-ide baru ke dalam bidang kecerdasan buatan:

- Mengembangkan algoritma baru. Algoritma adalah salah satu elemen dari kecerdasan buatan. Algoritma baru dapat digunakan untuk meningkatkan kinerja sistem kecerdasan buatan dalam berbagai tugas.
- Menemukan data baru. Data adalah bahan bakar bagi kecerdasan buatan. Data baru dapat digunakan untuk melatih model kecerdasan buatan yang lebih akurat dan dapat diandalkan.
- Menemukan cara baru untuk menerapkan kecerdasan buatan. Kecerdasan buatan dapat diterapkan dalam berbagai bidang, seperti kesehatan, manufaktur, dan transportasi. Ada banyak peluang untuk menemukan cara baru untuk menerapkan kecerdasan buatan untuk menyelesaikan persoalan dunia nyata.

Kecerdasan buatan saat ini mencakup berbagai macam subbidang, seperti pembelajaran mesin, logika dan inferensi, pemrosesan bahasa alami, pengenalan pola, serta robotika. Setiap subbidang memiliki fokus dan tantangannya sendiri. Dengan mempelajari berbagai subbidang kecerdasan buatan, seseorang dapat mengembangkan pemahaman yang lebih komprehensif tentang bidang ini dan mempersiapkan diri untuk karier yang sukses di masa depan.

# AI Prehistory

Philosophy	logic, methods of reasoning mind as physical system foundations of learning, language, rationality
Mathematics	formal representation and proof algorithms, computation, (un)decidability, (in)tractability probability
Psychology	adaptation phenomena of perception and motor control experimental techniques (psychophysics, etc.)
Economics	formal theory of rational decisions
Linguistics	knowledge representation grammar
Neuroscience	plastic physical substrate for mental activity
Control theory	homeostatic systems, stability simple optimal agent designs

S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



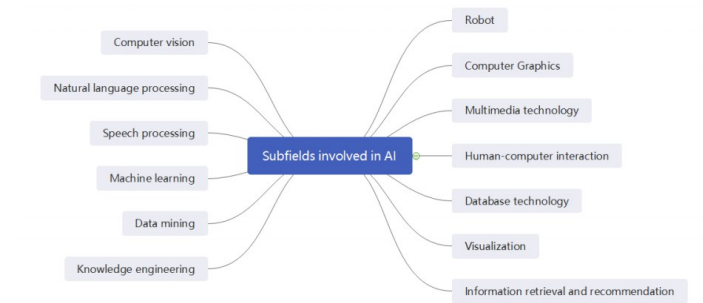
Sejarah kecerdasan buatan (AI) dapat ditelusuri jejaknya kembali ke awal abad ke-20, ketika para filsuf, matematikawan, psikolog, ekonom, ahli bahasa, neurosaintis, dan ahli teori kontrol mulai memikirkan cara membuat alat yang dapat melakukan tugas-tugas yang biasanya dilakukan oleh manusia.

- **Filsafat** telah memainkan peran penting dalam perkembangan AI. Filsuf telah merenungkan masalah seperti sifat kecerdasan, kesadaran, dan pikiran selama berabad-abad. Pemikiran mereka telah membantu para peneliti AI untuk memahami apa yang ingin mereka capai dan bagaimana mereka dapat mencapainya.
- **Matematika** juga merupakan dasar penting dari AI. Algoritma yang digunakan oleh sistem AI didasarkan pada matematika. Matematikawan telah mengembangkan teknik-teknik untuk membantu para peneliti AI untuk memecahkan masalah kompleks dan untuk membuat sistem AI yang lebih akurat dan efisien.
- **Psikologi** telah memberikan banyak wawasan tentang cara kerja kecerdasan manusia. Penelitian psikologi telah membantu para peneliti AI untuk memahami bagaimana manusia berpikir, belajar, dan menyelesaikan persoalan.
- **Ekonomi** telah memberikan wawasan tentang bagaimana AI dapat digunakan untuk meningkatkan efisiensi dan produktivitas. Ekonom telah membantu para peneliti AI untuk memahami bagaimana AI dapat digunakan untuk memecahkan

masalah dunia nyata, seperti pemborosan dan kemiskinan.

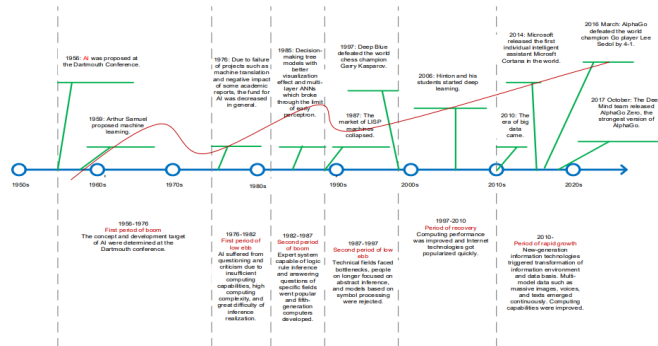
- **Linguistik** telah memberikan wawasan tentang cara kerja bahasa manusia. Penelitian linguistik telah membantu para peneliti AI untuk mengembangkan sistem yang dapat memahami dan menghasilkan bahasa manusia.
- **Neurosains** telah memberikan wawasan tentang cara kerja otak manusia. Penelitian neurosains telah membantu para peneliti AI untuk mengembangkan sistem yang dapat meniru cara kerja otak manusia.
- **Teori kontrol** telah memberikan wawasan tentang cara membuat sistem yang dapat beroperasi secara mandiri. Penelitian teori kontrol telah membantu para peneliti AI untuk mengembangkan sistem yang dapat belajar dari pengalaman dan menyesuaikan diri dengan lingkungan mereka.

# Sub-fields of AI



AI Development Report 2020

# Brief Development History



HUAWEI

1053001-AI | Institut Teknologi Del

6

- 1950: Alan Turing menerbitkan makalah "*Computing Machinery and Intelligence*", yang memperkenalkan tes Turing, sebuah tes untuk menilai kecerdasan mesin.
- 1956: John McCarthy, Marvin Minsky, Claude Shannon, dan Nathaniel Rochester mengadakan konferensi Dartmouth, yang dianggap sebagai kelahiran resmi AI.
- 1958: John McCarthy menciptakan istilah "Artificial Intelligence".
- 1959: Arthur Samuel mengembangkan program permainan catur pertama yang dapat belajar dari pengalamannya.
- 1966: Marvin Minsky dan Seymour Papert menerbitkan buku "Perception" yang memperkenalkan konsep jaringan saraf.
- 1970-an: Era AI yang suram, ketika banyak peneliti AI menjadi skeptis tentang kemungkinan menciptakan kecerdasan buatan yang sebenarnya.
- 1980-an: Kebangkitan AI, ketika peneliti AI mulai mengembangkan teknik baru yang lebih menjanjikan.
- 1987: Pengenalan pola wajah pertama kali digunakan oleh IBM untuk memverifikasi identitas pelanggan di ATM.
- 1990-an: Era AI yang eksplosif, ketika kemajuan dalam komputasi dan data membuat pengembangan AI lebih mudah.
- 1997: Deep Blue, komputer yang diprogram oleh IBM, mengalahkan Garry Kasparov, juara dunia catur saat itu.



- 2000-an: Era AI yang canggih, ketika AI digunakan dalam berbagai aplikasi, seperti robotika, perawatan kesehatan, dan transportasi.
- 2010-an: Era AI yang masif, ketika AI digunakan dalam berbagai aplikasi, seperti pengenalan suara, terjemahan bahasa, dan pembelajaran mesin.
- 2020-an: Era AI yang revolusioner, ketika AI digunakan untuk mengubah berbagai aspek kehidupan manusia.

# Four Views of AI

Systems that think like humans	Systems that think rationally
<p>"The exciting new effort to make computers think ... <i>machines with minds</i>, in the full and literal sense" (Haugeland, 1985)</p> <p>"The automation of activities that we associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)</p>	<p>"The study of mental faculties through the use of computational models" (Charniak and McDermott, 1985)</p> <p>"The study of the computations that make it possible to perceive, reason, and act" (Winston, 1992)</p>
Systems that act like humans	Systems that act rationally
<p>"The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990)</p> <p>"The study of how to make computers do things at which, at the moment, people are better" (Rich and Knight, 1991)</p>	<p>"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkoff, 1990)</p> <p>"The branch of computer science that is concerned with the automation of intelligent behavior" (Luger and Stubblefield, 1993)</p>

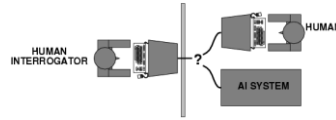
S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



Empat kemungkinan tujuan yang ingin dicapai dalam kecerdasan buatan (AI):

- **Sistem yang berpikir seperti manusia.** Tujuannya adalah untuk mengembangkan sistem AI yang dapat berpikir dan bernalar dengan cara yang sama seperti manusia. Hal ini melibatkan pengembangan sistem yang dapat memahami bahasa alami, belajar dari pengalaman, dan membuat keputusan berdasarkan informasi yang tidak lengkap.
- **Sistem yang bertindak seperti manusia.** Tujuannya adalah untuk mengembangkan sistem AI yang dapat berperilaku dengan cara yang tidak dapat dibedakan dengan manusia. Hal ini melibatkan pengembangan sistem yang dapat berinteraksi dengan dunia secara alami, dan dapat memahami serta merespons emosi manusia.
- **Sistem yang berpikir rasional.** Tujuannya adalah untuk mengembangkan sistem AI yang dapat berpikir logis dan mengambil keputusan rasional. Hal ini melibatkan pengembangan sistem yang dapat mengikuti aturan dan prosedur, serta dapat menghindari kesalahan.
- **Sistem yang bertindak rasional.** Tujuannya adalah untuk mengembangkan sistem AI yang dapat berperilaku bermanfaat bagi manusia. Hal ini mencakup pengembangan sistem yang dapat menghindari kerugian terhadap manusia, dan dapat meningkatkan kesejahteraan manusia.

# Acting Humanly: The Turing Test



- A problem questioned by AI researchers for years: “when can we count a **machine** as being **intelligent?**”
- The most famous response is attributed to Alan Turing, a British mathematician and computing pioneer. The famous “Turing Test”, based on ideas he expressed in a paper published in 1950 : “Computing Machinery and Intelligence”.
  - Human interrogates entity via teletype for 5 minutes. If, after 5 minutes, human cannot tell whether entity is human or machine, then the entity must be counted as intelligent.
- To date, no program has yet passed the Turing Test! However, there is the annual Loebner Prize which awards scientists for getting close. See [https://en.wikipedia.org/wiki/Loebner\\_Prize](https://en.wikipedia.org/wiki/Loebner_Prize) for more information.
- In order to pass the Turing Test, a program that succeeded would need to possess: **natural language processing**, **knowledge representation**, **automated reasoning**, and **machine learning**. → **basic Turing Test**

S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

8

Tes Turing adalah tes kemampuan mesin untuk menunjukkan perilaku cerdas yang setara, atau tidak dapat dibedakan, dengan perilaku manusia. Tes ini diperkenalkan oleh Alan Turing dalam makalahnya tahun 1950, "*Computing Machinery and Intelligence*".

Tes Turing dilakukan oleh interogator manusia yang terlibat dalam percakapan bahasa alami dengan dua pihak lain: manusia dan mesin yang dirancang untuk menghasilkan teks mirip manusia. Jika interogator tidak dapat membedakan mesin dari manusia dengan pasti, mesin tersebut dikatakan telah lulus ujian.

Tes Turing bukannya tanpa kritik. Ada yang berpendapat bahwa hal tersebut terlalu antroposentris, artinya bias terhadap kecerdasan manusia. Yang lain berpendapat bahwa ini terlalu mudah untuk dipertainkan, karena mesin dapat diprogram untuk memberikan respons spesifik yang dirancang untuk mengelabui interogator.

Meskipun memiliki keterbatasan, Tes Turing tetap menjadi tolok ukur penting dalam bidang kecerdasan buatan. Hal ini telah membantu mendorong penelitian terhadap pembelajaran mesin dan pemrosesan bahasa alami, dan telah menginspirasi pengembangan sistem AI baru yang mampu melakukan perilaku yang semakin mirip

manusia.

# Thinking Humanly: Cognitive Science

- Thinking humanly means trying to understand and model how the human mind works.
- There are (at least) two possible routes that humans use to find the answer to a question:
  1. We reason about it to find the answer. This is called “*introspection*”.
  2. We conduct *experiments* to find the answer, drawing upon scientific techniques to conduct controlled experiments and measure change.
- The field of *Cognitive Science* focuses on modeling how people think.

S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



# Thinking Rationally: Laws of Thought

- The “thinking rationally” approach to AI uses *symbolic logic* to capture the laws of rational thought as symbols that can be manipulated.
- *Reasoning* involves manipulating the symbols according to well-defined rules, kind of like algebra.
- The result is an idealized model of human reasoning. This approach is attractive to theoretists, i.e., modeling how humans should think and reason in an ideal world.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



# Acting Rationally

- **Rational** behavior: doing the **right thing**.
- The right thing: that which is expected to maximize goal achievement, given the available information.
- Doesn't necessarily involve thinking - e.g., blinking reflex - but thinking should be in the service of rational action.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



# What are Intelligences?

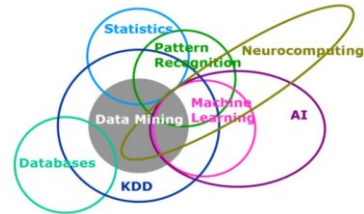
## HOWARD GARDNER'S THEORY OF MULTIPLE INTELLIGENCES





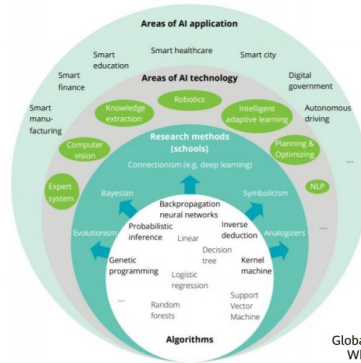
# What is AI?

- AI is a science that studies and develops theories, methods, techniques, and application systems for simulating and extending human intelligence.
- In 1956, John McCarthy proposed the concept of AI:
  - “science and engineering of making intelligent machines, especially intelligent computer program”.
- AI is concerned with making machines work in an intelligent way, similar to the way that the human mind works.



Identification of concepts related to AI and machine learning  
AI Development Report 2020

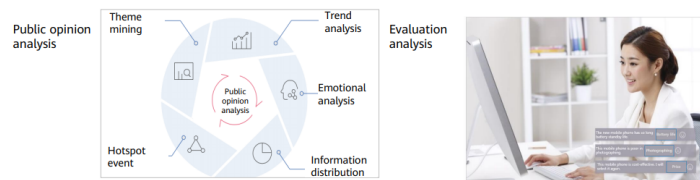
# Technical Fields and Application Fields of AI



Global AI Development  
White Paper 2020

# NLP Application Scenario (1)

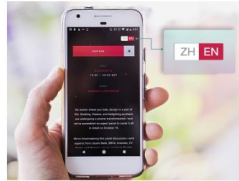
- The main topics of NLP research include machine translation, text mining, and sentiment analysis. NLP imposes high requirements on technologies but confronts low technology maturity. Due to high complexity of semantics, it is hard to reach the human understanding level using parallel computing based on big data and parallel computing only.
- In future, NLP will achieve more growth: understanding of shallow semantics → automatic extraction of features and understanding of deep semantics; single-purpose intelligence (ML) → hybrid intelligence (ML, DL, and RL)
- Application scenarios:



HUAWEI

## NLP Application Scenario (2)

Machine translation



Text classification



- Other applications:
  - Knowledge graph
  - Intelligent copywriting
  - Video subtitle
  - ...



# Our Goal

- Build rational agents.
  - What is Agent?
  - What is Rational?

# Agents and Environments

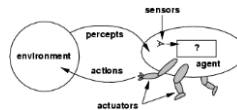
- An *agent* is an entity that *perceives* its environment through sensors and *acts* within its environment through actuators.

- Abstractly, an *agent* is a function from percept histories to actions:

$$f: P^* \rightarrow \mathcal{A}$$

For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance

- *Agents* include: humans, robots, software, etc. Sensors? Actuators?



S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Rationality

- What is “rational” depends on:
  1. Performance measures (how, when)
  2. The agents’ prior knowledge of the environment
  3. The actions the agent can perform
  4. Percept sequence to date (history): everything agent has perceived so far
  
- Rational  $\neq$  omniscient, clairvoyant
- Rationality maximizes expected performance
- Perfection maximizes actual performance
  
- Rational  $\Rightarrow$  exploration, learning, autonomy

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

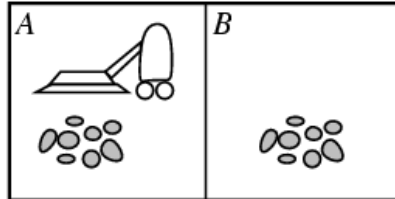
# Rational Agents

- A **rational agent** chooses whichever action maximizes the expected value of the performance measure given the percept sequence to date.
  - i.e., an agent that is designed to play a game should make moves that increase its chances of winning the game.
- Caveat: **computational limitations make perfect rationality unachievable**
  - design best **program** for given machine resources

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



# Vacuum-Cleaner World



- **Percepts:** location and contents, e.g., [A, Dirty]
- **Actions:** *Left, Right, Suck, NoOp*

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

21

Untuk mengilustrasikan gagasan ini, kita menggunakan contoh yang sangat sederhana – dunia penyedot debu. Dunia khusus ini hanya memiliki dua lokasi: kotak A dan B. Agen vakum mendeteksi di kotak mana ia berada dan apakah ada kotoran di dalam kotak tersebut.

Ia bisa memilih bergerak ke kiri, ke kanan, menyedot kotoran, atau tidak sama sekali.

# A Vacuum-Cleaner Agent

Percept sequence	Action
[A, Clean]	Right
[A, Dirty]	Suck
[A, Clean]	Left
[A, Dirty]	Suck
[A, Clean], [A, Clean]	Right
[A, Clean], [A, Dirty]	Suck
<b>function</b> REFLEX-VACUUM-AGENT( <i>location, status</i> ) <b>returns</b> an action	
<b>if</b> <i>status</i> = <i>Dirty</i> <b>then return</b> <i>Suck</i>	
<b>else if</b> <i>location</i> = <i>A</i> <b>then return</b> <i>Right</i>	
<b>else if</b> <i>location</i> = <i>B</i> <b>then return</b> <i>Left</i>	

- What is the **right** function?
- Can it be implemented in a small agent program?

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

22

Salah satu fungsi agen yang sangat sederhana adalah sebagai berikut: jika kotak saat ini kotor, maka menyedot; jika tidak, lanjutkan ke kotak lainnya.

Namun, hal ini menimbulkan pertanyaan: Bagaimana cara mengisi tabel yang benar? Dengan kata lain, apa yang membuat suatu agen menjadi baik atau buruk, cerdas atau bodoh?

# Performance Measure

- Fixed performance measure evaluates the environment sequence
  - one point per square cleaned up in time  $t$ ?
  - one point per clean square per time step, minus one per move?
  - penalize for  $> k$  dirty squares?

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

23

# PEAS

- To design a rational agent, we must specify the **task environment**
  - Performance measure??
  - Environment??
  - Actuators??
  - Sensors??
- Consider, e.g., the task of designing an automated taxi.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

24

Kinerja – kualitas apa yang harus dimiliki agen?

Lingkungan – dimana agen harus bertindak?

Aktuator – bagaimana cara agen melakukan tindakan?

Sensor – bagaimana agen mempersepsikan lingkungan?

# PEAS: Automated Taxi

- **Performance measure??** safety, destination, profits, legality, comfort, ...
- **Environment??** US streets/freeways, traffic, pedestrians, weather, ...
- **Actuators??** steering, accelerator, brake, horn, speaker/display, ...
- **Sensors??** video, accelerometers, gauges, engine sensors, keyboard, GPS, ...

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

25

# PEAS: Internet Shopping Agent

- **Performance measure??** price, quality, appropriateness, efficiency
- **Environment??** current and future WWW sites, vendors, shippers
- **Actuators??** display to user, follow URL, fill in form
- **Sensors??** HTML pages (text, graphics, scripts)

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Environment Types

- **Fully observable** (vs. partially observable):
  - An agent's sensors give it access to the complete state of the environment at each point in time.
- **Deterministic** (vs. stochastic):
  - The next state of the environment is completely determined by the current state and the action executed by the agent. (If the environment is deterministic except for the actions of other agents, then the environment is strategic)
- **Episodic** (vs. sequential):
  - The agent's experience is divided into atomic "episodes" (each episode consists of the agent perceiving and then performing a single action), and the choice of action in each episode depends only on the episode itself.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Environment Types

- **Static** (vs. dynamic):
  - The environment is unchanged while an agent is deliberating. (The environment is **semi-dynamic** if the environment itself does not change with the passage of time but the agent's performance score does)
- **Discrete** (vs. continuous):
  - A limited number of distinct, clearly defined percepts and actions.
- **Single agent** (vs. multi-agent):
  - An agent operating by itself in an environment.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del



# Agent Types

- Four basic types in order of increasing generality:
  - simple reflex agents
  - reflex agents with state
  - goal-based agents
  - utility-based agents
- All these can be turned into learning agents

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Environment Types

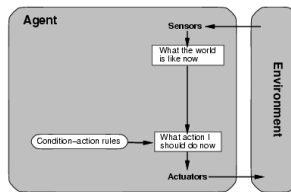
	Solitaire	Backgammon	Internet shopping	Taxi
Observable??	Yes	Yes	No	No
Deterministic??	Yes	No	Partly	No
Episodic??	No	No	No	No
Static??	Yes	Semi	Semi	No
Discrete??	Yes	Yes	Yes	No
Single-agent??	Yes	No	Yes (except auctions)	No

- The environment type largely determines the agent design
- The real world is (of course) partially observable, stochastic, sequential, dynamic, continuous, multi-agent

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Simple reflex agents

- Simple look-up table, mapping percepts to actions, is out of question (too large, too expensive to build)
- Many situations can be summarized by condition-action rules (humans: learned responses, innate reflexes)



Rectangles: agent's internal state

Implementation: easy

Ovals: background information

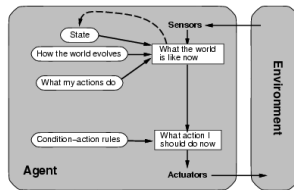
Applicability: narrow

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.

1053001-AI | Institut Teknologi Del

# Simple reflex agents with state

- Sensory information alone is not sufficient
- Need to keep track of how the world evolves  
(evolution: independently of agent, or caused by agent's actions)

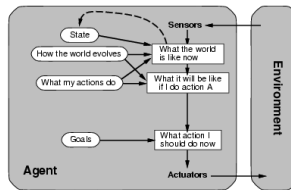


- How the world evolved: model-based agent

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Goal-based agents

- State & actions don't tell where to go
- Need goals to build sequences of actions (planning)
- Goal-based: uses the same rules for different goals



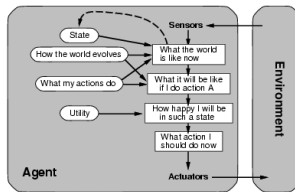
- Reflex: will need a complete set of rules for each goal

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.

1053001-AI | Institut Teknologi Del

# Utility-based agents

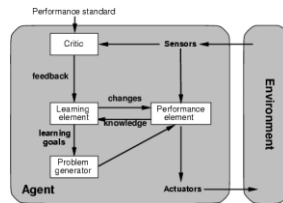
- Several action sequences to achieve some goal (binary process)
- Need to select among actions & sequences. Preferences.
- Utility: State  $\rightarrow$  real number (express degree of satisfaction, specify trade-offs between conflicting goal)



S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.  
1053001-AI | Institut Teknologi Del

# Learning agents

- Agent operates in an initially unknown environment, and becomes more competent than its initial knowledge alone might allow
- Learning: process of modification of each component of the agent to bring the components into closer agreement with the available



- feedback information, thus improving overall performance of the agent.

S. J. Russell and P. Borvig, *Artificial Intelligence: A Modern Approach (4th Edition)*, Prentice Hall International, 2021.

1053001-AI | Institut Teknologi Del

eof

10S3001-AI | Institut Teknologi Del