

Politechnika Poznańska

**Adam Meissner**

Adam.Meissner@put.poznan.pl

<http://www.man.poznan.pl/~ameis>

**SZTUCZNA INTELIGENCJA (SI, AI)**

**Wprowadzenie do sztucznej inteligencji**

# Literatura (1)

## Podstawowa

- [1] Flasiński M., *Wstęp do sztucznej inteligencji*, PWN, 2011.
- [2] Russell S.J., Norvig P., *Artificial Intelligence: A Modern Approach, Third Edition*, Prentice Hall, 2010.

## Uzupełniająca

- [3] Bostrom N., *Superinteligencja: scenariusze, strategie, zagrożenia*, Wyd. Helion, 2016.
- [4] Kurzweil R., *Nadchodzi osobliwość: kiedy człowiek przekroczy granice biologii*, Wyd. Kurhaus, 2013.
- [5] Labenz P., *Konwersacyjna sztuczna inteligencja: od Turinga do Potiomkina*, w: *kognitywistyka: o umyśle umyślnie i nieumyślnie*, 2004
- [6] LindenberG G., *Ludzkość poprawiona. Jak najbliższe lata zmienią świat, w którym żyjemy*, Wyd. Otwarte, 2018.

## Literatura (2)

### Uzupełniająca (cd.)

- [7] Nilsson N.J., *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann Pub., 1998.
- [8] Turing A., *Computing Machinery and Intelligence*, *Mind* 49: 433-460, 1950.

# SI – co to znaczy?

**Znaczenie terminu *sztuczny*:** „przymiotnik określający substancje, przedmioty lub zjawiska stworzone przez człowieka w celu zastąpienia odpowiedników naturalnych” [za: *Słownik języka polskiego*, PWN, 2019].

**Znaczenie terminu *inteligencja*:**

- „Zdolność do postrzegania, wnioskowania, uczenia się, porozumiewania się i działania w złożonym środowisku” [7]
- trudności ze sformułowaniem definicji trafnej i pełnej
- „Jedno z najbardziej wieloznacznych pojęć w psychologii odnoszące się do sprawności w zakresie czynności poznawczych” [Encyklopedia PWN, 2019]

**Klasyfikacja przybliżonych definicji SI (wg [2])**

Działanie jak człowiek	Działanie racjonalne
Myślenie jak człowiek	Myślenie racjonalne

# SI – działanie jak człowiek

- rola wizjonerskiego artykułu A. Turinga pt. *Computing Machinery and Intelligence* z 1950 roku [8]
- **test Turinga** – skonfrontowanie badanego obiektu z komisją ekspertów
- kluczowe elementy systemu podchodzącego do testu Turinga:
  - analizator języka naturalnego
  - odpowiednio reprezentowana wiedza przedmiotowa
  - mechanizm wnioskujący
  - mechanizm uczenia się
  - system wizyjny i manipulatory (w rozszerzonej wersji testu)
- problem: test Turinga jest bardzo trudny do opisania formalnego
- zaskoczenie – program ELIZA (Weizenbaum, 1966) protoplasta współczesnych chatbotów [5].

# SI – myślenie jak człowiek

- powiązania z badaniami z zakresu **kognitywistyki** dotyczącymi modelowania ludzkiego mózgu i umysłu
- system *General Problem Solver* (A. Newell & H.A. Simon, 1957)
  - początkowo przeznaczony do dowodzenia prostych twierdzeń matematycznych, rozwijany jako uniwersalna maszyna do rozwiązywania problemów „dobrze zdefiniowanych”
  - wiedza w systemie obejmowała rozpatrywany problem, warunki rozwiązania problemu, a także obiekty składowe i operacje na nich
  - system wykonywał operacje na obiektach, poszukując kolejnych pośrednich (częściowych) rozwiązań problemu o malejącej odległości do rozwiązania końcowego
- architektury SOAR, GOMS, KLM – modele (i ich implementacje) inspirowane działaniem ludzkiego umysłu, w szczególności opisujące zjawiska, które występują podczas interakcji człowiek-komputer

# SI – myślenie racjonalne

- logika jako fundament myślenia racjonalnego
- rachunek sylogizmów Arystotelesa (IV w. p.n.e)
  - sylogizm jest to schemat wnioskowania, składający się z dwóch przesłanek oraz wniosku
  - przesłanki zawierają element wspólny, a każdy element wniosku zawarty jest w dokładnie jednej przesłance
  - przykład (tzw. sylogizm Arystotelesa)

*Każdy M jest P*

*Każdy S jest M*

---

*Każdy S jest P*

- zamysł sformułowania uniwersalnego formalizmu logicznego *lingua characteristica*, G. Leibniz (XVII w.)
- systemy dowodzenia twierdzeń, systemy wnioskujące

# SI – działanie racjonalne

- działanie racjonalne rozumie się jako dążenie do osiągnięcia założonego celu (lub maksymalne zbliżenie się do niego) z wykorzystaniem całej dostępnej informacji
- często jest wspierane racjonalnym myśleniem; nie jest to jednak konieczne — np. prawidłowe działanie odruchowe
- jednym ze sposobów modelowania działań racjonalnych jest ujęcie agentowe [2]
- podstawowe cechy agenta
  - dobrze zdefiniowana granica między agentem a otoczeniem
  - autonomia (zdolność do podejmowania samodzielnych decyzji i działań)
  - interaktywność (zdolność do komunikowania się z otoczeniem, w tym z innymi agentami)
  - reaktywność (zdolność do reagowania na zmiany w otoczeniu, np. spowodowane działaniem innych agentów)



# SI – klasyfikacja podejść

- jednym z kryteriów klasyfikowania podejść (ujęć) wyodrębnianych w dziedzinie SI jest sposób reprezentowania i przetwarzania wiedzy; z tej perspektywy można wyróżnić *ujęcie symboliczne* oraz *ujęcie subsymboliczne* [7]
- **ujęcie symboliczne:** wiedza jest reprezentowana w postaci *napisów*, czyli ciągów znaków (symboli); przetwarzanie wiedzy polega na działaniu na napisach
- **ujęcie niesymboliczne:** wszelkie inne formy reprezentowania wiedzy np. liczby, wektory, sygnały w sztucznych sieciach neuronowych

# SI – rys historyczny (1)

- IV w. p.n.e., Arystoteles – wnioskowanie za pomocą sylogizmów
- XIV w., R. Llull – dzieło *Ars Magna*, „generator idei”
- przełom XVI i XVII w., rabin Jehuda Loew ben Bazalel z Pragi – Golem
- XVII w., G. Leibniz – *lingua characteristica*
- XVIII w., A. Smith – podstawy ekonometrii
- druga połowa XIX w., G. Boole, G. Frege – podstawy logiki pierwszego rzędu
- pierwsza połowa XX w., D. Hilbert, K. Goedel, A. Church, A. Turing, T. Skolem, J. Herbrand – badania logiką pierwszego rzędu, sformułowanie teorii wnioskowania automatycznego
- rok 1943, W. McCulloch i W. Pitts – proste modele neuronów, system SNARC (M. Minsky i D. Edmonds, 1950)
- lata 40-te XX w., A. Turing, J. von Neumann, J. Mauchly, J. Atanasoff, K. Zuse – pierwsze maszyny liczące (Colossus, Eniac, ABC, Z3)

# SI – rys historyczny (2)

- lata 50-te:
  - pierwsze „silniki” gier dwuosobowych, np. program A. Samuela grający w warcaby
  - pierwsze systemy uczenia maszynowego (A. Samuel)
  - warstwowe sztuczne sieci neuronowe, np. Percptron (F. Rosenblatt), Adaline/Madaline (B. Widrow)
- rok 1956, konferencja w Dartmouth – termin *sztuczna inteligencja*
- okres wielkich oczekiwań (1956-1975)
  - programy do dowodzenia twierdzeń, np. GPS, LT (A. Newell, H.A. Simon)
  - język LISP, system *Advice Taker* (J. McCarthy, 1958)
  - sformułowanie reguły (wnioskowania) rezolucji (J.A. Robinson, 1965)
  - system MACSYMA (C. Engelman, 1968)
  - język Prolog (A. Colmerauer, R. Kowalski, 1971)
  - system SHRDLU (T. Winograd, 1972)
  - raport Lighthilla (1973)

## SI – rys historyczny (3)

- początki systemów z wiedzą (1970-1980)
  - DENDRAL (B. Buchanan et al., 1965)
  - MYCIN (E. Feigenbaum, 1976)
- okres dojrzały (od 1980)
  - zastosowanie systemów eksperckich w przemyśle (np. PROSPECTOR, G2)
  - „zmartwychwstanie” sztucznych sieci neuronowych
  - projekt „maszyn piątej generacji”
  - wpływ SI na inne nauki: logika formalna (np. logiki deskrypcyjne), probablistyka (np. formalizm sieci Bayesa)

# SI – współczesność

- przez współczesność rozumie się okres po 2010 roku
- sukces paradygmatu ***deep learning***
- oprogramowanie komputera Watson IBM (np. gracz w *Jeopardy!*, 2011)
- samochód autonomiczny, np. *Google Car*, 2012
- model/system *Word2Vec*, T. Mikolov et al. (Google), 2013
- program *AlphaGo*, DeepMind (Google), 2015
- program *Watson Debater*, IBM, 2018

# SI – przykłady zastosowań [6]

- gra w różne gry (np. szachy, Go, poker) na poziomie nie ustępującym człowiekowi lub wyższym, np. program *AlphaGo Zero*, DeepMind, 2017
- rozpoznawanie mowy ze skutecznością przewyższającą ekspertów (ludzi), 2017
- rozpoznawanie obrazów, np. twarzy – projekt Bystre Oko, Chiny (w toku)
- stawianie diagnoz medycznych z trafnością (w niektórych dziedzinach) przewyższającą doświadczonych lekarzy
- profilowanie klientów (np. w bankowości)

# Plan kolejnych wykładów

1. Podstawy logiki pierwszego rzędu
2. Modelowanie problemów za pomocą grafu stanów
3. Problemy spełniania ograniczeń i metody ich rozwiązywania
4. Gry dwuosobowe z pełną informacją o sumie zerowej
5. Elementy wnioskowania automatycznego
6. Elementy logik deskrypcyjnych
7. Architektury systemów eksperckich, systemy regułowe
8. Reprezentowanie i przetwarzanie wiedzy o czasie
9. Elementy uczenia maszynowego
10. Podstawy sztucznych sieci neuronowych
11. Zastosowania SI we technice współczesnej