

Spis treści

<i>Przedmowa</i>	11
<i>Podziękowania</i>	15
1. Dotychczasowe dokonania i obecne możliwości	17
Fazy wzrostu w dziejach ludzkości	17
Wielkie oczekiwania	20
Okresy nadziei i rozczarowań	23
Aktualny stan badań	32
Opinie na temat przyszłości sztucznej inteligencji	38
2. Ścieżki wiodące ku superinteligencji	45
Sztuczna inteligencja	46
Emulacja mózgu	56
Poznanie biologiczne	64
Interfejsy mózg – komputer	76
Sieci i organizacje	81
Podsumowanie	83
3. Formy superinteligencji	87
Superinteligencja szybka	88
Superinteligencja zbiorowa	89
Superinteligencja jakościowa	93
Osiągalność pośrednia i bezpośrednia	95
Źródła przewagi inteligencji cyfrowej	96

4. Dynamika eksplozji inteligencji	101
Moment i prędkość odejścia	101
Oporność	106
Siła optymalizacyjna i eksplozywność	117
5. Decydująca przewaga strategiczna	123
Czy wiodący projekt uzyska decydującą przewagę strategiczną?	124
Jak duży będzie projekt uwieńczony sukcesem?	129
Od decydującej przewagi strategicznej do singletonu	134
6. Poznawcze supermoce	139
Funkcjonalności i supermoce	140
Scenariusz przejścia władzy przez SI	144
Władza nad naturą i bytami posiadającymi moc sprawczą	148
7. Pobudki superinteligencji	159
Związek pomiędzy inteligencją a motywacją	159
Konwergencja instrumentalna	164
8. Czy czeka nas zagłada?	173
Zagłada ludzkości jako najbardziej prawdopodobny skutek eksplozji inteligencji?	173
Zdradziecki zwrot	175
Rodzaje złośliwych usterek	179
9. Problem kontroli	191
Dwa problemy agencji	191
Metody kontroli potencjału	194
Metody doboru motywacji	205
Podsumowanie	213
10. Wyrocznie, dziny, suwereni i narzędzia	215
Wyrocznie	215
Dziny i suwereni	219
Narzędziowa SI	223
Porównanie	230

11. Scenariusze wielobiegunowości	233
O koniach i ludziach	234
Życie w gospodarce algorytmicznej	243
Poprzejsiowie tworzenie się singletonu?	258
12. Zaszczepianie wartości	271
Problem przekazywania wartości	271
Selekcja ewolucyjna	274
Uczenie ze wzmocnieniem	275
Stopniowy przyrost wartości	276
Zrąb systemu wartości	279
Uczenie się wartości	281
Modyfikowanie emulacji	292
Projektowanie organizacji instytucjonalnej	294
Podsumowanie	300
13. Wybór kryteriów wyboru	303
Potrzeba normatywności pośredniej	303
Spójna, ekstrapolowana wola	306
Modele moralności	315
Zrób to, co mam na myśli	318
Lista komponentów	320
Dostateczne przybliżenie	328
14. Perspektywa strategiczna	331
Strategia naukowo-technologiczna	332
Scenariusze i czynniki umożliwiające ich realizację	347
Współpraca	356
15. Moment krytyczny	369
Filozofia z nieprzekraczalnym terminem	369
Co mamy do zrobienia?	371
Proszę o powstanie najlepsze cechy natury ludzkiej	375
<i>Przypisy</i>	377
<i>Bibliografia</i>	455
<i>Skorowidz</i>	483