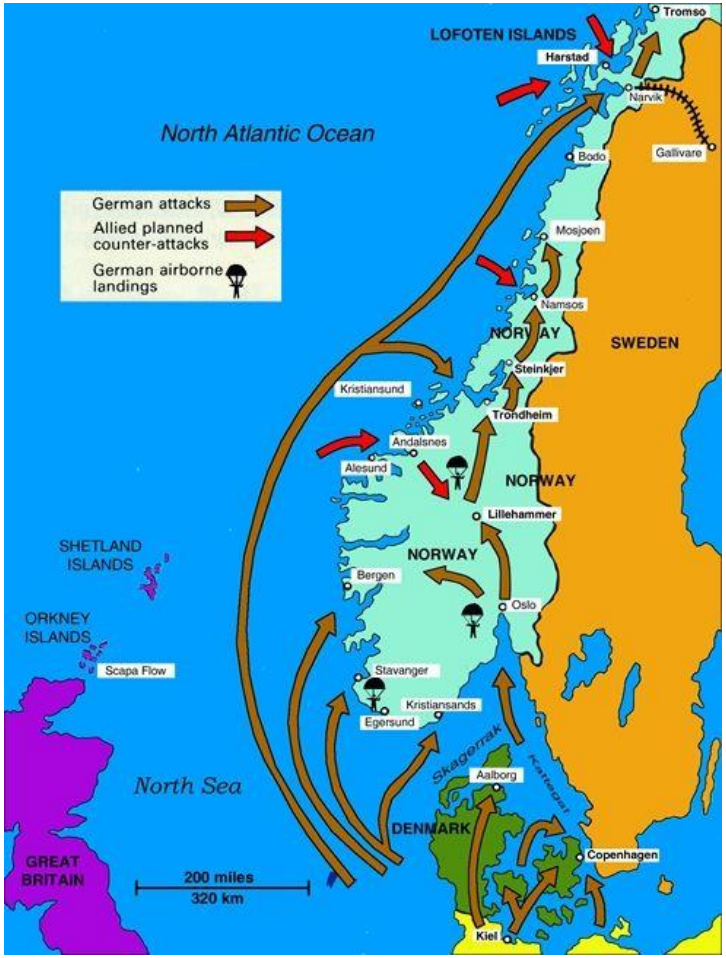
Australian Government
Department of Defence
Defence Science and Technology Group

Scenarios for Trusted Autonomous Systems

Brandon Pincombe
Head Land Organisational and Management Science,
Land Capability Analysis Branch
Joint and Operational Analysis Division
Defence Science and Technology Group

Norway: April and June 1940

- 9 April 1940: Germany invaded Norway
  - Rapidly consolidated hold on south.
  - British TF landed in central Norway from 12 April; force out by 2 May.
  - Franco-Polish-British TF landed in north from 14 April; supporting Norwegian 6th Division; made gains, but …
  - 10 May Germany invaded France, Belgium and Netherlands; allied priorities changed;
  - Recapture Narvik 28 May but Norwegian royal family and cabinet evacuated on 7 June & allied forces followed on 8 June;
  - HMS Glorious was on her fourth mission to northern Norway;
  - German SAG had left Kiel 4 June to strike allied positions in northern Norway.



8 June 1940: Norwegian Sea

Outline

- The future is uncertain, so use preparatory scenarios to assess the adaptability of options.
  - This can be applied to autonomous systems.
- Have used experts to identify critical variables and build scenarios within the dimensional space created by these.
  - Experientially based – captures variables that *have* been critical.
  - Used to test adaptability; some ways of inculcating adaptability may
    - Extend outside the ranges considered on variables; or (rarely?)
    - Extend to other variables
- Open question: autonomous generation of scenarios?

The Future is *Uncertain*

… the major cultural barrier to scenario implementation stems from the way we define managerial competence. Good managers, we say, *know* where they are, where they're going, and how they'll get there. *We equate managerial competence with "knowing,"* and assume that decisions depend on facts about the present and about the future. Of course, the reality is that *we have no facts about the future*.

-- Ian Wilson (2000) "From Scenario Thinking to Strategic Action" Technological Forecasting and Social Change, 65(1) 23-29.

Some managers think that it just takes the right experts to be able to predict the future … but:

- "The phonograph … is not of any commercial value"
  Thomas Edison (inventor of the phonograph) c. 1880
- "Heavier than air flying machines are impossible."
  Lord Kelvin (Royal Society president) c.1895
- "I think there is a world market for about five computers."
  Thomas J. Watson (Chairman of IBM) 1943
- "There is no reason for any individual to have a computer in their home."
  Ken Olson (Chairman of DEC) 1977

May in Tanunda	
Hottest Ever	28.9°
Average Max.	17.5°
Average Min.	6.9°
Coldest Ever	-2.1°
Av. Rainy Days	7.3

Plasticity only exists within bounds …

Venus	
Hottest Ever	462°
Coldest Ever	462°

- **Autonomy**: ability of an organism or system, within its social context, to operate under unpredictable future environmental states.
  - **Plasticity**: the adaptability of an organism to changes in its environment or differences between its various habitats.

Australia	
Hottest Ever	50.7°
Coldest Ever	-23.0°

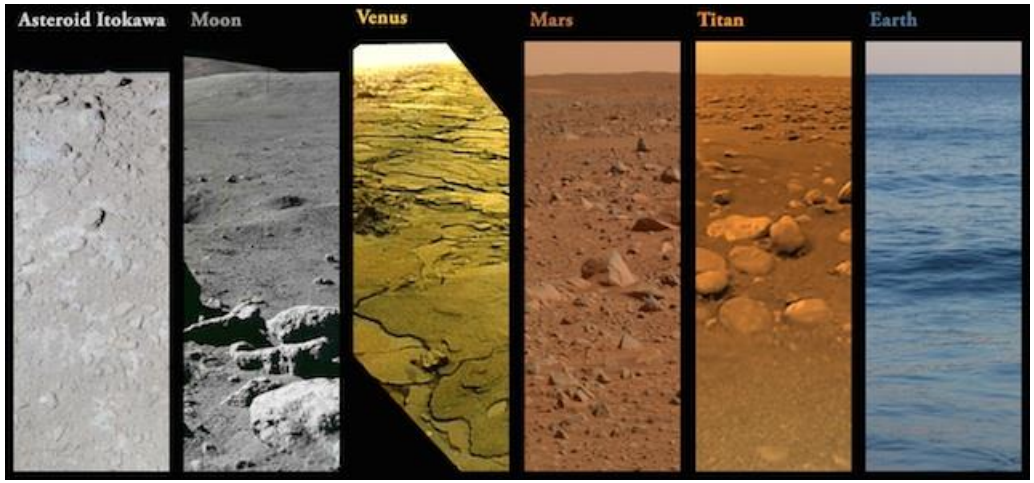
Earth	
Hottest Ever	56.7°
Coldest Ever	-89.2°

Mercury	
Hottest Ever	695°
Coldest Ever	-220°

Mars	
Hottest Ever	20°
Coldest Ever	-153°

May in Tanunda	
Hottest Ever	28.9°
Average Max.	17.5°
Average Min.	6.9°
Coldest Ever	-2.1°
Av. Rainy Days	7.3

- What are the minimum bounds on the autonomous capabilities we need in certain circumstances?
- How do we identify, evaluate and prioritise autonomous capabilities?
- What are the maximum bounds on the autonomy of various systems?



Predictive vs. Preparatory Scenarios

- Preparatory scenarios require us to identify the *critical uncertainties* … these define the bounds of the space considered (& are the dimensions of the scenario space).
- Then scenarios can be developed by selecting points or regions within the scenario space … or scenarios can be eschewed and dimensions used directly.
- **Robust** options deal with all scenarios; **Plastic** ones can adapt to all scenarios.
- First … find the dimensions.

Predictive Scenarios

- Each situation that needs to be dealt with has one scenario.
  - May need many scenarios.
  - Solutions are predicted futures.
  - Brittle if predictions wrong.
  - Enemies tend to change themselves to make your predictions wrong.
  - Even experts have futurology problems.

Preparatory Scenarios

- Focus on covering the critical uncertainties (the *dimensions*).
  - Shell's pre-oil-shock scenarios didn't *predict* an oil embargo but *prepared* for a medium term supply squeeze.
  - Divide into 'intuitive logic' and 'indeterminist' approaches.
  - Right critical uncertainties vital!

- Pincombe, Blunden, Pincombe & Dexter (2012) "Ascertaining a Hierarchy of Dimensions from Time-Poor Experts: linking tactical vignettes to strategic scenarios" *Technological Forecasting and Social Change* http://dx.doi.org/10.1016/j.techfore.2012.05.001
- Wright &Cairns (2011) *Scenario Thinking: Practical Approaches to the Future*. Houndmills, UK: Palgrave Macmillan.
- Derbyshire & Wright (2014) Preparing for the future: Development of an 'antifragile' methodology that complements scenario planning by omitting causation. *Technological Forecasting and Social Change, 82, pp. 215-225* doi:10.1016/j.techfore.2013.07.001.
- Giffin & Reid (2003) "A Woven Web of Guesses, Canto Two: Network Centric Warfare and the Myth of Inductivism" http://www.dodccrp.org/events/8th_ICCRTS/pdf/109.pdf

Delphi-like process for ascertaining critical variables

- The fundamentals of the Delphi process are:
  - anonymity throughout
  - individual contributions to an initial round of questions
  - assessment of responses
  - controlled feedback to participants with revised questions
  - opportunity to comment, revise/amend contributions based on the feedback
  - start out qualitatively and end up quantitatively
  - cycle through until consensus (usually at least 3 surveys)
  - arrive at consensus or otherwise based on statistical sampling

- Dalkey & Helmer (1951) The Use of Experts for the Estimation of Bombing Requirements. RM727
- Jablin & Sussman (1978) An exploration of communication and productivity in real brainstorming groups, *Hum. Commun. Res.* 4 329–337.
- Pincombe, Blunden, Pincombe & Dexter (2012) "Ascertaining a Hierarchy of Dimensions from Time-Poor Experts: linking tactical vignettes to strategic scenarios" *Technological Forecasting and Social Change* http://dx.doi.org/10.1016/j.techfore.2012.05.001

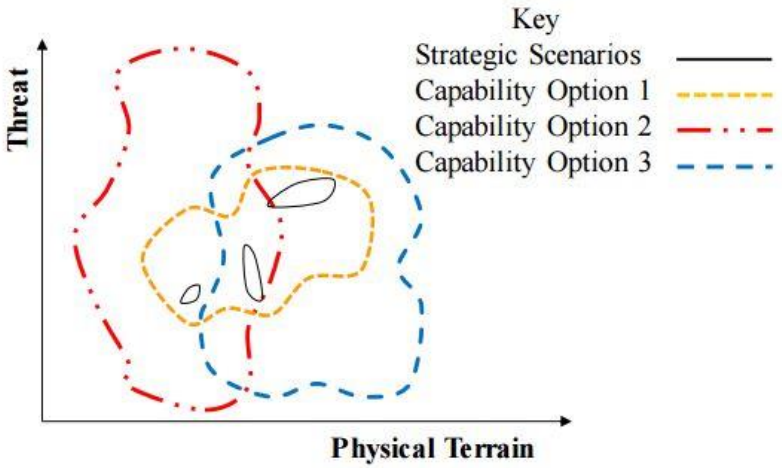
Experts for dimensions of future land warfare contexts for the Australian Army …

- Obviously who we ask is **very** important:
  - Some 8-year old boys I polled had definite ideas on future land warfare:
    - very large nuclear powered tanks with rotating mincing flails;
    - giant flying androids dispensing *viridian* green poison gas;
    - genetically engineered trained paranoid *Canadian* arboreal octopi; …
  - Maharishi Mahesh Yogi reckons that yogic flying is the way to go. The NLP's 1997 platform promised an army of 7,000 yogic flyers and thought our main threat was evil sentient Antarctic penguins.
- We need military experts … for this work we asked:
  - 93 experienced and educated military experts **drawn from all corps** [as well as 25 ACSC students].
  - Representation of all stakeholder groups is ***very*** important.
- In the process of ascertaining the dimensions for the use of autonomous systems in HADR operations

• Pincombe, Blunden, Pincombe & Dexter (2012) "Ascertaining a Hierarchy of Dimensions from Time-Poor Experts: linking tactical vignettes to strategic scenarios" *Technological Forecasting and Social Change* http://dx.doi.org/10.1016/j.techfore.2012.05.001

Computational Scenarios & Feasible Scenario Space

- The scenario space is a set of variables that can be fed into a multi-variate function:
  - For software autonomous agents: they are multi-variate functions that could be tested directly;
  - For physical autonomous agents: a model could be established and tested.
    - Errors could be introduced in the modelling process.
- FSS: a surface which covers the set of scenario parameters for which a given capability set can achieve success, within acceptable levels of inherent risk.
  - About robustness rather than plasticity.
- Not identifying a critical variable produces big problems for both.



- On-going CS work: Hussein Abbass -- postgrad project in "computational scenario planning for disaster management" [see: https://www.unsw.adfa.edu.au/computational-scenario-planning-disaster-management].

Baker, Bender, Abbass & Sarker "A Scenario-based Evolutionary Scheduling Approach for Assessing Future Supply Chain Capabilities" pp. 485-511 in Evolutionary Learning Vol 49 of Studies in Computational Intelligence. Springer.
Abbass, Bender, Dam, Baker, Whitacre & Sarker (2008) "Computational scenario based capability planning" GECCO'08 1437-1444. 10.1145/1389095.1389378
Abbass, Bender, Gaidow & Whitbread (2011) "Computational Red Teaming: Past, Present and Future" IEEE Computational Intelligence, 6(1) 30-42. 10.1109/MCI.2010.939578
Bow den, Pincombe & Williams (2015) "Feasible Scenario Spaces: a new w ay of measuring capability impacts" http://www.mssanz.org.au/modsim2015/D3/bow den.pdf

Autonomous discovery of dimensions?

- The present method of identifying dimensions is based on past experience
- For the dimensions we have we know that tests of robustness and plasticity are okay, just like regression is okay, within the range that we have samples for.
  - Outside this range we hope that options are still plastic on identified critical variables
- However, as with the example of HMS Glorious, the biggest problem is likely to come from critical variables that have not been identified.
- Open question: is it possible to develop an autonomous system to identify unknown critical variables.
  - May be based on trying to produce shocks in the options/strategies/agents (or models of them) operating within the scenario space.

Questions?

Inductive vs. Indeterminist scenarios

Intuitive Logic

- Emphasises the causal unfolding of future events
- Inductive
- Can give a misleading impression as to the usefulness of 'weak signals' or 'early warnings'

Indeterminist

- Views uncertainty as originating from indeterminism
- Critical rationalist

- Bradfield, Derbyshire & Wright (2016) "Augmenting the intuitive logics scenario planning method for a more comprehensive analysis of causation" *International Journal of Forecasting* 10.1016/j.ijforecast.2016.01.004.
- Cairns, Wright & Fairbrother (2016) "Promoting articulated action from diverse stakeholders in response to public policy scenarios: A case analysis of the use of 'scenario improvisation' method"*Technological Forecasting and Social Change*, **103**, 97-108 10.1016/j.techfore.2015.10.009.
- Bradfield, Derbyshire & Wright (2016) 'The critical role of history in scenario thinking: Augmenting causal analysis within the intuitive logics scenario development methodology" Futures 10.1016/j.futures.2016.02.002.
- Derbyshire & Wright (2014) Preparing for the future: Development of an 'antifragile' methodology that complements scenario planning by omitting causation. *Technological Forecasting and Social Change,* **82**, pp. 215-225 doi:10.1016/j.techfore.2013.07.001.
- Rhyne (1974) "Technological forecasting within alternative whole futures projections" *Technological Forecasting and Social Change*,**6**:133-162.

Wisdom of the crowds?

Polgar, Nakamura, Vachier-Lagrave + "the world" voting on which of their moves to adopt.

Carlsen – coherent plan.

Sources of failure

- Most errors come from ignoring or not seeking disconfirming information:
  - Traffic analysis showing Kreigsmarine SAG (ignoring);
  - Lack of air patrol or crows nest watcher (not seeking);
- It often isn't that we don't have sufficient information; it is that we aren't asking the right questions (and, maybe, aren't asking questions at all).
- Darryn will probably say:
  - Put most of your resources into hedging against unacceptable outcomes through cautious investments and constant search for disconfirmation;
  - Put what you can afford to lose into options to take advantage of opportunities and search for confirming evidence that you haven't failed yet.

Degeneracy

Whitacre & Bender (2010) "Degeneracy: a design principle for achieving robustness and evolvability " Journal of Theoretical Biology, 263(1) 143-153.
https://arxiv.org/ftp/arxiv/papers/0907/0907.0510.pdf
Whitacre (2010) "Degeneracy: a link betw een evolvability, robustness and complexity in biological systems" Theoretical Biology and Medical Modelling, 7(6) 10.1186/1742-4682-7-6

Evolution takes ages …

- "analyse[ing] genome-wide data from 51 Eurasians from ~45,000–7,000 years ago … the proportion of Neanderthal DNA decreased from 3–6% to around 2%, consistent with natural selection against Neanderthal variants in modern humans." – Fu (and 63 others!!) (2016) "The genetic history of Ice Age Europe" Nature doi:10.1038/nature17993

- ***38,000 years*** to halve Neanderthal DNA due to selection/drift.