An agent-based iterated learning model for understanding language evolution

George Sains

george.sains@bristol.ac.uk

28th October 2022

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.
  - Language exists as **utterances** of pairs of meanings and signals.

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.
  - Language exists as **utterances** of pairs of meanings and signals.
  - Signals are understood using a neural network (taught with back-propagation) and meanings are encoded using the **Obverter procedure** [Oliphant and Batali, 1997].

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.
  - Language exists as **utterances** of pairs of meanings and signals.
  - Signals are understood using a neural network (taught with back-propagation) and meanings are encoded using the **Obverter procedure** [Oliphant and Batali, 1997].
- Main measurements are:

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.
  - Language exists as **utterances** of pairs of meanings and signals.
  - Signals are understood using a neural network (taught with back-propagation) and meanings are encoded using the **Obverter procedure** [Oliphant and Batali, 1997].
- Main measurements are:
  - **Expressibility**: The portion of the meaning space that can be represented by an individuals language.

The Iterated Learning Model [Kirby and Hurford, 2002]

- Language evolution model with a focus on the **Critical Period Hypothesis**.
- Consists of a **Meaning & Signal space**, which are an abstract idea of language represented with binary strings.
  - Language exists as **utterances** of pairs of meanings and signals.
  - Signals are understood using a neural network (taught with back-propagation) and meanings are encoded using the **Obverter procedure** [Oliphant and Batali, 1997].
- Main measurements are:
  - **Expressibility**: The portion of the meaning space that can be represented by an individuals language.
  - **Stability**: How similar two agents languages are.

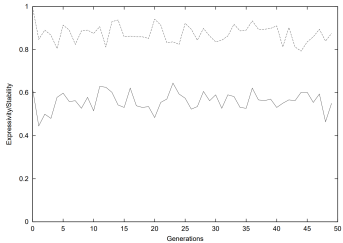
Obverter Procedure

α	β	γ	δ	R
0.2	0.25	0.1	0.3	a
0.2	0.25	0.1	0.4	b
0.4	0.25	0.1	0.3	c
0.2	0.25	0.7	0.0	d

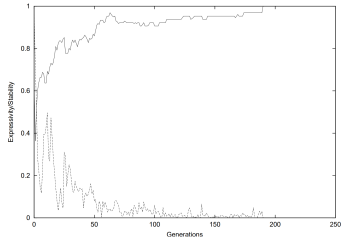
$\Rightarrow$

S	α	β	γ	δ
a	0.0	0.0	0.0	1.0
b	0.0	0.0	0.0	1.0
c	1.0	0.0	0.0	0.0
d	0.0	0.0	1.0	0.0

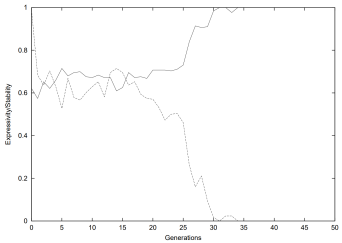
ILM Behaviour



(20 Utterances per generation)



(2000 utterances per generation)



(50 utterances per generation)

Expansion to Agent-Based Model

- **Goal**: A model for language evolution which simulates known ideas in linguistics.

Expansion to Agent-Based Model

- **Goal**: A model for language evolution which simulates known ideas in linguistics.
- Place agents on a (*fully-connected*) Erdős–Rényi graph who periodically communicate with neighbours.

Expansion to Agent-Based Model

- **Goal**: A model for language evolution which simulates known ideas in linguistics.
- Place agents on a (*fully-connected*) Erdős–Rényi graph who periodically communicate with neighbours.
- Agents have two states, child and adult. At end of childhood perform Obvert procedure and allow agent to communicate.

Expansion to Agent-Based Model

- **Goal**: A model for language evolution which simulates known ideas in linguistics.
- Place agents on a (*fully-connected*) Erdős–Rényi graph who periodically communicate with neighbours.
- Agents have two states, child and adult. At end of childhood perform Obvert procedure and allow agent to communicate.
- Agents have an increasing chance to die with age and be replaced with a new child agent.

Expansion to Agent-Based Model

- **Goal**: A model for language evolution which simulates known ideas in linguistics.
- Place agents on a (*fully-connected*) Erdős–Rényi graph who periodically communicate with neighbours.
- Agents have two states, child and adult. At end of childhood perform Obvert procedure and allow agent to communicate.
- Agents have an increasing chance to die with age and be replaced with a new child agent.
- Child agents learn in the same way as the ILM, adults still use back-propagation but will only **partial-obvert**.

Expansion to Agent-Based Model (Partial Obvert)

α	β	γ	δ	R
0.2	0.25	0.1	0.3	a
0.2	0.25	0.1	0.4	b
0.4	0.25	0.1	0.3	c
0.2	0.25	0.7	0.0	d

$\Rightarrow$

S	α	β	γ	δ
a	0.0	0.0	0.0	1.0
b	0.0	0.0	0.0	1.0
c	1.0	0.0	0.0	0.0
d	0.0	0.0	1.0	0.0

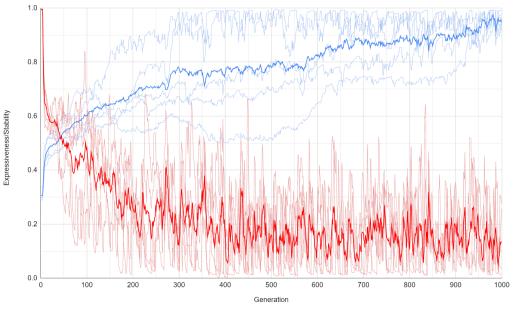
Receive (δ,c)

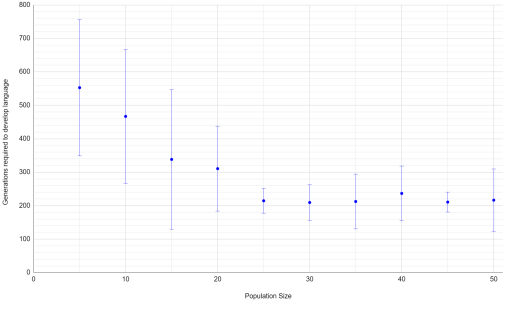
α	β	γ	δ	R
0.2	0.25	0.1	0.2	a
0.2	0.25	0.1	0.2	b
0.4	0.25	0.1	**0.6**	c
0.2	0.25	0.7	0.0	d

$\Rightarrow$

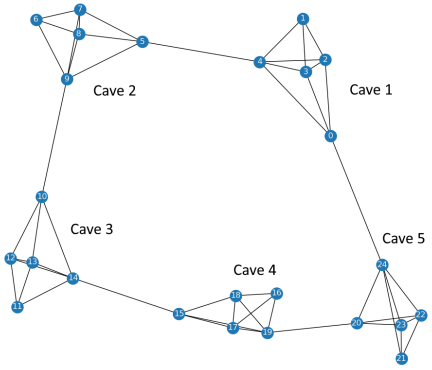
S	α	β	γ	δ
a	0.0	0.0	0.0	*1.0*
b	0.0	0.0	0.0	*1.0*
c	0.0	0.0	0.0	**1.0**
d	0.0	0.0	1.0	0.0

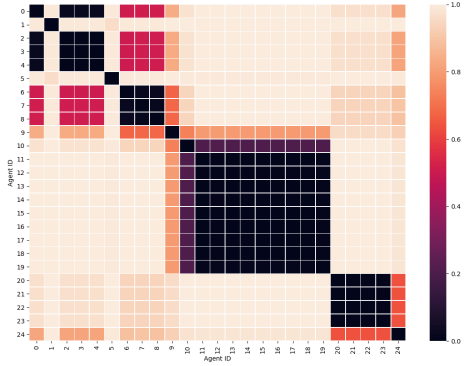
Results


Expressiveness/Stability
Generation
0.0
0.2
0.4
0.6
0.8
1.0
0
100
200
300
400
500
600
700
800
900
1000


Generations required to develop language
Population Size
0
100
200
300
400
500
600
700
800
0
10
20
30
40
50

Small-World Theory [Telesford et al., 2011]





Future Work

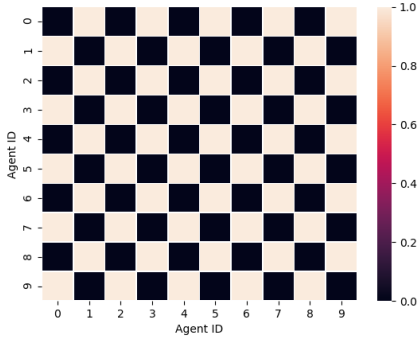
- Alternative to the Obvert procedure as it's **computationally costly**.

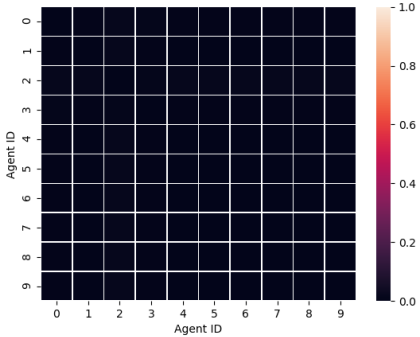
Future Work

- Alternative to the Obvert procedure as it's **computationally costly**.
- Investigate **Creolization** as this behaviour isn't present in this model [Hymes, 1971].

Future Work

- Alternative to the Obvert procedure as it's **computationally costly**.
- Investigate **Creolization** as this behaviour isn't present in this model [Hymes, 1971].





Thank you! Any questions?

Bibliography

[Hymes, 1971] Hymes, D. H. (1971).
Pidginization and creolization of languages.
CUP Archive.

[Kirby and Hurford, 2002] Kirby, S. and Hurford, J. R. (2002).
The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model, pages 121–147.
Springer London, London.

[Oliphant and Batali, 1997] Oliphant, M. and Batali, J. (1997).
Learning and the emergence of coordinated communication.
Center for Research on Language Newsletter, 11.

[Telesford et al., 2011] Telesford, Q. K., Joyce, K. E., Hayasaka, S., Burdette, J. H., and Laurienti, P. J. (2011).
The ubiquity of small-world networks.
Brain connectivity, 1(5):367–375.