Bargaining over a common categorisation

Marco LiCalzi & Nadia Maagli

VOLUME 193 No. 3 March 2016

SYNTHESE

AN INTERNATIONAL JOURNAL

FOR EPISTEMOLOGY, METHODOLOGY

AND PHILOSOPHY OF SCIENCE

Editors-in-Chief:
Otávio Bueno, Wiebe van der Hoek & Gila Sher

SPECIAL ISSUE GUEST EDITORS:

Thomas Ågotnes, Giacomo Bonanno and Wiebe van der Hoek

Special Issue on LOGIC AND THE FOUNDATIONS OF GAME
AND DECISION THEORY (LOFT2014)

🦄 Springer

ISSN 0039–7857

🦄 Springer

Springer

Bargaining over a common categorisation

Marco LiCalzi[1] · **Nadia Maagli**[2]

Abstract Two agents endowed with different categorisations engage in bargaining to reach an understanding and agree on a common categorisation. We model the process as a simple non-cooperative game and demonstrate three results. When the initial disagreement is focused, the bargaining process has a zero-sum structure. When the disagreement is widespread, the zero-sum structure disappears and the unique equilibrium requires a retraction of consensus: two agents who individually associate a region with the same category end up rebranding it under a different category. Finally, we show that this last equilibrium outcome is Pareto dominated by a cooperative solution that avoids retraction; that is, the unique equilibrium agreement may be inefficient.

1 Introduction

It is widely documented that agents organise information by means of categories, with significant implications over their behaviour (Cohen and Lefebvre 2005). This paper is a theoretical foray in a strictly related but still poorly explored territory: what kind

✉ Marco LiCalzi
  licalzi@unive.it

  Nadia Maagli
  nadia.maagli@malix.univ-paris1.fr

[1] Department of Management, Università Ca' Foscari Venezia,
  S. Giobbe, Cannaregio 873, 30121 Venice, Italy

[2] EDE-EM, Université Paris 1-Panthéon-Sorbonne, Paris, France

of outcome may emerge when two agents endowed with individual categorisations interact and develop a common categorisation?

There exist different families of models for categorical reasoning; see Sect. 1 in Kruschke (2008) for a concise overview. The model developed in this paper borrows from the theory of conceptual spaces, proposed in Gärdenfors (2000) as an alternative approach for the modelling of cognitive representations. A tenet of this theory is the claim that natural concepts may be associated with convex regions of a suitable space and, in particular, that a *conceptual space* consists of a collection of convex regions. This underlying geometric structure resonates with early theories of categorisation based on prototypes (Rotsch 1975; Mervis and Rotsch 1981), and has recently been given both evolutionary (Jäger 2007) and game-theoretic foundations (Jäger et al. 2011).

Conceptual spaces, on the other hand, provide a representational framework that may accommodate different notions. Recently, Gärdenfors (2014) has expanded their scope towards semantics and the study of meaning. In particular, Warglien and Gärdenfors (2013) suggest an interpretation of semantics as a mapping between individual conceptual spaces. People negotiate meaning by finding ways to map their own personal categorisations to a common one; see Warglien and Gärdenfors (2015) for an insightful discussion with references. A well-known example is the integration of different cultures within an organisation, when different communication codes blend into a commonly understood language (Wernerfelt 2004).

Warglien and Gärdenfors (2013) rely on the theory of fixed points to argue for the plausibility of two individuals achieving a "meeting of minds" and sharing a common conceptual space. Their approach, however, is merely existential and thus offers no insight in the structure of the possible outcomes associated with establishing a common conceptual space. We shed a constructive light by framing the problem of how two agents reach a common understanding as the equilibrium outcome of a bargaining procedure.

We borrow from the theory of conceptual spaces the assumption that agents' categorisations correspond to a collection of convex categories or, for short, to a *convex categorisation*. However, the neutrality of this latter term is meant to help the reader keeping in mind that our results are consistent with, but logically independent from, the theory of conceptual spaces.

We analyse a simple non-cooperative game where two agents, endowed with their own individual convex categorisations, negotiate over the construction of a common convex categorisation. Agents exhibit stubbornness as they are reluctant to give up on their own categorisation, but they are engaged in a dialectic process that must ultimately lead to a common categorisation. The common convex categorisation emerges as the (unique) equilibrium of the game, aligning it with the argument that meaning is constructed and shared via an equilibrating process (Parikh 2010).

We demonstrate two main phenomena, depending on whether the disagreement between agents' individual spaces is focused or widespread. Under focused disagreement, the bargaining process has a zero-sum structure: agents' stubbornness leads to a unique equilibrium where each concedes as little as possible, and the agents who has a larger span of control over the process ends up being better off. Under widespread disagreement, the zero-sum structure disappears and each agent confronts a dilemma:

Fig. 1 A binary convex
categorisation
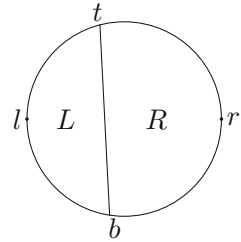
holding on to one of his individual categories weakens his position on another one. At the unique equilibrium, these conflicting pressures force a retraction of consensus: two agents who individually agree on a region falling under the same category end up relabeling it in order to minimise conflict. Moreover, we uncover that convex categorisations may be a source of inefficiency: the equilibrium outcome is Pareto dominated by the Nash bargaining solution without retraction.

The rest of the paper is organized as follows. Section 2 describes our game-theoretic model. Section 3 defines two forms of disagreement (focused and widespread) and states our results as theorems. Section 4 provides concluding comments. All proofs are relegated in the appendix.

2 Model

There are two agents. Each agent $i = 1, 2$ has his own binary convex categorisation over the closed unit disk C in $\mathbb{R}^2$. Our qualitative results carry through for any convex compact region C in $\mathbb{R}^2$, but this specific choice is elegant and analytically advantageous because C is invariant to rotations. Interestingly, Jäger and Van Rooij (2007) also choose to develop their second case study under the assumption that the meaning space is circular. Conventionally, we label the two concepts L for Left and R for Right and use them accordingly in our figures.

The agents agree on the classification of two antipodal points in C: they both label $l = (-1, 0)$ as L and $r = (0, 1)$ as R, respectively. Intuitively, this implies that the agents' categorisations are not incompatible. More formally, define the intersection of the agents' initial categorisations as their *shared (partial) categorisation*. Under our assumption, the individual categorisations are compatible because the shared categorisation is not empty. On the other hand, since in general the individual categorisations are different, the shared categorisation is only partial. The agents' problem is to move from their (partial) shared categorisation to a common (total) categorisation.

The categorisation of Agent i over C consists of two convex regions L_i and R_i. Dropping subscripts for simplicity, this may look like in Fig. 1. Clearly, the representation is fully characterized by the chord $\overline{tb}$ separating the two convex regions. The endpoints t and b for the chord are located in the top and in the bottom semicircumference, respectively. (To avoid trivialities, assume that the antipodal points l and r are interior.) The two convex regions of the categorisation may differ in extension and thus the dividing chord need not be a diameter for C.
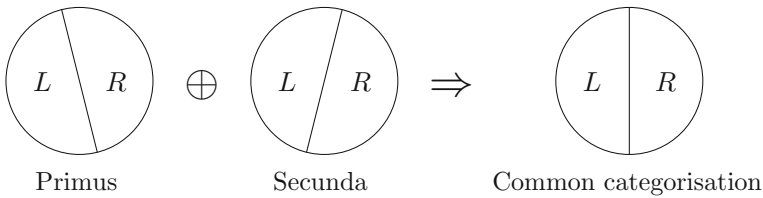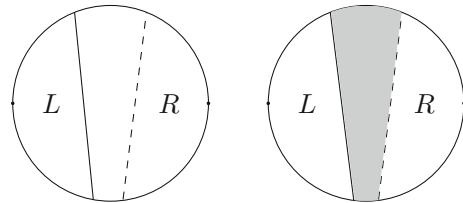
Fig. 2 The search for a common categorisation

Fig. 3 The disagreement area

Consider the categorisations of the two agents. Unless $\overline{t_1 b_1} = \overline{t_2 b_2}$, the regions representing the concepts are different and the shared categorisation is partial. If the agents are to reach a common categorisation, they must negotiate an agreement and go through a bargaining process over categorisations, where each agent presumably tries to push for preserving as much as possible of his own original individual categorisation. Figure 2 provides a pictorial representation for the process: Agent 1 (Primus) and Agent 2 (Secunda) negotiate a common categorisation as a compromise between their own individual systems of categories.

We provide a simple game–theoretic model for their interaction and study the equilibrium outcomes. We do not claim any generality for our model, but its simplicity should help making the robustness of our results transparent.

The two agents play a game with complete information, where the endpoints (t_i, b_i) of each agent i are commonly known. Without any loss of generality, let Primus be the agent for whom t_1 precedes t_2 in the clockwise order. Primus picks a point t in the arc interval $[t_1, t_2]$ from the top semicircumference, while Secunda simultaneously chooses a point b between b_1 and b_2 from the bottom semicircumference. The resulting chord $\overline{tb}$ defines the common categorisation. Under our assumption that the antipodal points l and r are interior, the agents cannot pick either of them.

Each agent evaluates the common categorisation against his own. Superimposing these two spaces, there is one region where the common categorisation and the individual one agree and (possibly) a second region where they disagree. For instance, consider the left-hand side of Fig. 3 where the solid and the dashed chords represent the agent's and the common categorisation, respectively. The two classifications disagree over the central region, coloured in grey on the right-hand side.

Each agent wants to minimise the disagreement between his own individual and the common categorisation. For simplicity, assume that the payoff for an agent is the opposite of the area of the disagreement region D; that is, $u_i = -\lambda(D_i)$ where λ is the Lebesgue measure. (Our qualitative results carry through for any absolutely continuous measure μ.) Note that the region D need not be convex: when the chords

Fig. 4 Focused (*left*) and
widespread disagreement (*right*)
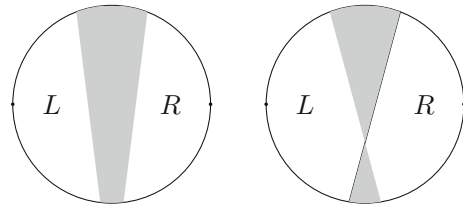
underlying the agent's and the common categorisation intersect inside the disc, D
consists of two opposing circular sectors.

3 Results

The study of the equilibria is greatly facilitated if we distinguish three cases. First,
when $t_1 = t_2$ and $b_1 = b_2$, the two individual categorisations (as well as the initial
shared categorisation) are identical: the unique Nash equilibrium has $t^* = t_1$ and
$b^* = b_2$, and the common categorisation agrees with the individual ones. This is a
trivial case, which we consider no further. From now on, we assume that the two
individual categorisations disagree and thus the initial shared categorisation is only
partial; that is, either $t_1 \neq t_2$ or $b_1 \neq b_2$ (or both).

The other two cases depend on the shape of the disagreement region D. When $\overline{t_1 b_1}$
and $\overline{t_2 b_2}$ do not cross inside the disc, then D is a convex set as in the left-hand side of
Fig. 4. We define this situation as *focused* disagreement, because one agent labels D as
L and the other as R. The disagreement is focused on whether D should be construed
as L or R.

Instead, when $\overline{t_1 b_1}$ and $\overline{t_2 b_2}$ cross strictly inside the disc, then D is the union of
two circular sectors as in the right-hand side of Fig. 4. This is the case of *widespread*
disagreement, because the two agents label the two sectors in opposite ways: the top
sector is L for one and R for the other, while the opposite holds for the bottom sector.

3.1 Focused disagreement

Under focused disagreement, t_1 precedes t_2 and b_2 precedes b_1 in the clockwise order.
The disagreement region is convex and the interaction is a game of conflict: as Primus's
choice of t moves clockwise, his disagreement region (with respect to the common
categorisation) increases, while Secunda's decreases. In particular, under our simpli-
fying assumption that payoffs are the opposites of the disagreement areas, this is a
zero-sum game.

Intuitively, players have opposing interests over giving up on their individual cate-
gorisations. Therefore, we expect that in equilibrium each player concedes as little as
possible. In our model, this leads to the stark result that they make no concessions at all
over whatever is under their control. That is, they exhibit maximal *stubbornness*. This
is made precise in the following theorem, that characterises the unique equilibrium.
All proofs are relegated in the appendix.

Fig. 5 The unique equilibrium
outcome under focused
disagreement
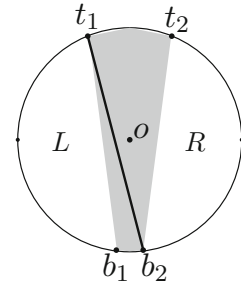
Theorem 1 *Under focused disagreement, the unique Nash equilibrium is $(t^*, b^*) = (t_1, b_2)$. Moreover, the equilibrium strategies are dominant.*

Figure 5 illustrates the equilibrium outcome corresponding to the situation depicted on the left-hand side of Fig. 4. The thick line defines the common categorisation. In this example, Primus and Secunda give up the small grey area on the left and on the right of the thick line, respectively. Note how Primus and Secunda stubbornly stick to their own original t_1 and b_2. Moreover, Primus gives up a smaller area and thus ends up being better off than Secunda. This shows that, in spite of its simplicity, the game is not symmetric. Our next result elucidates which player has the upper hand in general. Formally, let (t^s, b^s) be the Nash bargaining solution, with t^s and b^s being the midpoints of the two players' strategy sets. We say that in equilibrium Primus is *stronger* than Secunda if $u_1(t^*, b^*) \geq u_1(t^s, b^s) = u_2(t^s, b^s) \geq u_2(t^*, b^*)$.

To gain intuition, consider again Fig. 5. The thick line defining the common categorisation divides the disagreement region into two sectors $S_1(t_1 t_2 b_2)$ and $S_2(b_2 b_1 t_1)$. Primus wins S_1 and loses S_2; so he is stronger when $\lambda(S_1) \geq \lambda(S_2)$. The area of S_1 depends on the angular distance $\tau = \widehat{t_1 o t_2}$ controlled by Primus and on the angular distance $\theta_R = \widehat{t_2 o b_2}$ underlying the arc that is commonly labeled R; similarly, the area of S_2 depends on $\beta = \widehat{b_1 o b_2}$ and $\theta_L = \widehat{t_1 o b_1}$. Primus is advantaged when $\tau \geq \beta$ and $\theta_R \geq \theta_L$. The first inequality implies that his span of control is higher. The second inequality makes the common categorisation for R less contestable than for L, so that Primus' stubborn clinging to t_1 is more effective than Secunda's choice of b_2. The next result assumes that a player (say, Primus) has the larger span of control: then Primus is stronger when his span of control is sufficiently large, or when R is more contestable than L but the opponent's span of control is small enough.

Proposition 2 *Suppose $\tau \geq \beta$. If $\tau \geq \beta + (\theta_L - \theta_R)$, then Primus is stronger. If $\tau < \beta + (\theta_L - \theta_R)$, then there exists $\bar{\beta}$ such that Primus is stronger if and only if $\beta \leq \bar{\beta}$.*

3.2 Widespread disagreement

Under widespread disagreement, t_1 precedes t_2 and b_1 precedes b_2 in the clockwise order. The disagreement region is not convex and the interaction is no longer a zero-sum game. We simplify the analysis by making the assumption that the two chords
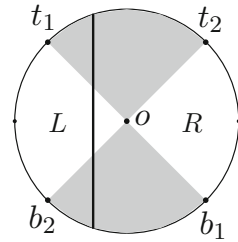
characterising the players' categorisations are diameters. Then the two angular dis-
tances $\tau = \widehat{t_1 o t_2}$ and $\beta = \widehat{b_1 o b_2}$ are equal, the players have the same strength and the
game is symmetric.

Players' stubbornness now has a double-edged effect, leading to a retraction of
consensus at the unique equilibrium. Before stating it formally, we illustrate this result
with the help of Fig. 6, drawn for the special case $\tau = \beta = \pi/2$. The thick line depicts
the common categorisation at the unique equilibrium for this situation.

Consider Primus. Choosing t very close to t_1 concedes little on the upper circular
sector, but exposes him to the risk of a substantial loss in the lower sector. This
temperates Primus' stubbornness and, in equilibrium, leads him to choose a value of
t^* away from t_1. However, as his opponent's choice makes the loss from the lower
sector smaller than the advantage gained in the upper sector, the best reply t^* stays
closer to t_1 than to t_2. An analogous argument holds for Secunda.

A surprising side-effect of these tensions is that, in equilibrium, the common cat-
egorisation labels the small white triangle between the thick line and the origin as
R, in spite of both agents classifying it as L in their own individual systems of cat-
egories. That is, in order to reach an agreement, players retract their consensus on
a small region and agree to recategorize part of their initial shared categorisation.
The following theorem characterise the unique equilibrium by means of the two
angular distances $\widehat{t^* o t_1}$ and $\widehat{b^* o b_2}$. It is an immediate corollary that the retraction
of consensus always occurs, unless $\tau = 0$ and the two agents start off with identical
categorisations.

Theorem 3 *Suppose that the individual categorisations are supported by diameters,
so that $\tau = \beta$. Under widespread disagreement, there is a unique Nash equilibrium
(t^*, b^*) characterised by*

$$\widehat{t^* o t_1} = \widehat{b^* o b_2} = \arctan \left(\frac{\sin \tau}{\sqrt{2} + 1 + \cos \tau} \right).$$

As the equilibrium necessitates a retraction of consensus, it should not be surprising
that we have an efficiency loss that we call the *cost of consensus*. The equilibrium
strategies lead to payoffs that are Pareto dominated by those obtained under different
strategy profiles. The following result exemplifies the existence of such cost using the
natural benchmark provided by the Nash bargaining solution (t^s, b^s), with t^s and b^s
being the midpoints of the respective arc intervals.

Proposition 4 *Suppose that the individual categorisations are supported by diameters. Under widespread disagreement, $u_i(t^*, b^*) \leq u_i(t^s, b^s)$ for each player $i = 1, 2$, with the strict inequality holding unless $\tau = 0$.*

4 Concluding comments

The game-theoretic model presented and solved in this paper is a mathematically reduced form, consistent with different interpretations. As discussed in the introduction, our motivation originates with a few recent contributions about the negotiation of meaning. Accordingly, we suggest to interpret the convex regions of a conceptual space as the (simplified) representation of lexical meanings for words (Gärdenfors 2014a). Each agent enters the negotiation with his own mapping between words and their meaning, and the purpose of their interaction is to generate a common mapping. This is a first step in the ambitious program of "modelling communication between agents that have different conceptual models of their current context", as proposed by Honkela et al. (2008).

If one also accepts the classical view that concepts have definitional structures, it is possibile to expand the scope of our model to the negotiation of concepts. However, we believe that the underlying philosophical difficulties make this a slippery path and we prefer to confine our discussion to the negotiation of lexical meaning for words. This places our contribution within the recent literature emphasising a game–theoretic approach to the analysis of language (Benz et al. 2005; Clark 2012; Parikh 2010).

Finally, we mention some advantages and limitations in our model. The use of noncooperative game theory highlights the "mixed motives" described in Warglien and Gärdenfors (2015): the negotiation agents have a common interest in achieving coordination on a common categorisation, tempered by individual reluctance in giving up their own categories. This conflict is a channel through which egocentrism affects pragmatics (Keysar 2007), and we show that it may impair efficiency. On the other hand, the simplicity of our model leaves aside important issues of context, vagueness and dynamics in the negotiation of the lexicon (Ludlow 2014).

Appendix: Proofs

Proof of Theorem 1

The proof is a bit long, but straightforward. It is convenient to introduce some additional notation. The endpoints (t_i, b_i) for the two agents' chords and their choices for t and
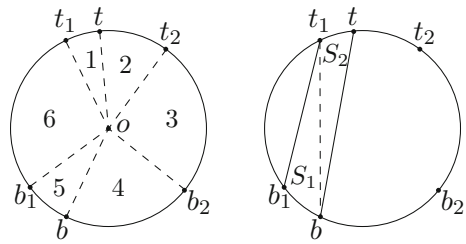
b identify six sectors. Proceeding clockwise, these are numbered from 1 to 6 on the
left-hand side of Fig. 7. For each sector i, we denote its central angle by θ_i; that is,
we let $\theta_1 = \widehat{t_1ot}$, $\theta_2 = \widehat{tot_2}$, $\theta_3 = \widehat{t_2ob_2}$, $\theta_4 = \widehat{b_2ob}$, $\theta_5 = \widehat{bob_1}$, and $\theta_6 = \widehat{b_1ot_1}$. The
following lemma characterises the disagreement area of each player as a function of
the six central angles.

Lemma 1 *The disagreement areas for Primus and Secunda are, respectively:*

$$\lambda(D_1) = \frac{\theta_1 + \theta_5 + \sin\theta_6 - \sin(\theta_1 + \theta_5 + \theta_6)}{2}, \tag{1}$$

and

$$\lambda(D_2) = \frac{\theta_2 + \theta_4 + \sin\theta_3 - \sin(\theta_2 + \theta_3 + \theta_4)}{2}.$$

Proof The disagreement region D_1 for Primus can be decomposed into the two sector-
like regions $S_1(t_1bb_1)$ and $S_2(t_1tb)$ as shown on the right-hand side of Fig. 7. (The
figure illustrates a special case, but the formulas hold in general.) We compute the
areas $\lambda(S_1)$ and $\lambda(S_2)$, and then add them up to obtain $\lambda(D_1)$.

Consider $S_1(t_1bb_1)$. It can be decomposed into two regions: the circular segment
from b to b_1 with central angle θ_5, and the triangle $T(t_1bb_1)$. The area of a circular
segment with central angle θ and radius r is $r^2(\theta - \sin\theta)/2$, which in our case reduces
to $(\theta_5 - \sin\theta_5)/2$. Concerning the triangle, the inscribed angle theorem implies that
the angle $\widehat{b_1t_1b} = \theta_5/2$; hence, by the law of sines, its area can be written as

$$\frac{\overline{t_1b} \times \overline{t_1b_1} \times \sin(\theta_5/2)}{2}. \tag{2}$$

Finally, by elementary trigonometry, $\overline{t_1b} = 2\sin[(\theta_5 + \theta_6)/2]$ and $\overline{t_1b_1} = 2\sin[(\theta_6)/2]$. Substituting into (2) and adding up the areas of the two regions, we obtain

$$\lambda(S_1) = \frac{\theta_5 - \sin\theta_5}{2} + 2\sin\left(\frac{\theta_5}{2}\right)\sin\left(\frac{\theta_6}{2}\right)\sin\left(\frac{\theta_5 + \theta_6}{2}\right).$$

By a similar argument, we obtain

$$\lambda(S_2) = \frac{\theta_1 - \sin\theta_1}{2} + 2\sin\left(\frac{\theta_1}{2}\right)\sin\left(\frac{\theta_5 + \theta_6}{2}\right)\sin\left(\frac{\theta_1 + \theta_5 + \theta_6}{2}\right).$$

Summing up $\lambda(S_1)$ and $\lambda(S_2)$, we find

$$\lambda(D_1) = \frac{\theta_1 - \sin\theta_1}{2} + \frac{\theta_5 - \sin\theta_5}{2}$$
$$+ 2\sin\left(\frac{\theta_5 + \theta_6}{2}\right)\left[\sin\left(\frac{\theta_1}{2}\right)\sin\left(\frac{\theta_1 + \theta_5 + \theta_6}{2}\right) + \sin\left(\frac{\theta_5}{2}\right)\sin\left(\frac{\theta_6}{2}\right)\right].$$

(3)

After some manipulations shown separately in the following Lemma 2, this expression simplifies to

$$\lambda(D_1) = \frac{\theta_1 + \theta_5 + \sin\theta_6 - \sin(\theta_1 + \theta_5 + \theta_6)}{2}.$$

The derivation of a specular formula for $\lambda(D_2)$ is analogous.   □

Lemma 2 *The expression in (3) for $\lambda(D_1)$ can be rewritten as*

$$\lambda(D_1) = \frac{\theta_1 + \theta_5 + \sin\theta_6 - \sin(\theta_1 + \theta_5 + \theta_6)}{2}.$$

Proof Let $p = \theta_5/2$ and $q = \theta_6/2$. Then

$$\lambda(S_1) = \frac{2p - \sin(2p)}{2} + 2\sin(p)\sin(q)\sin(p+q)$$
$$= \frac{2p - \sin(2p)}{2} + 2\sin(p+q)\left[\cos(p-q) - \cos(p+q)\right]$$
$$= \frac{2p - \sin(2p)}{2} + 2\sin(p+q)\cos(p-q) - \frac{\sin\left[2(p+q)\right]}{2}$$
$$= \frac{2p - \sin(2p)}{2} + \frac{\sin(2p) + \sin(2q)}{2} - \frac{\sin\left[2(p+q)\right]}{2}$$
$$= \frac{2p + \sin(2q) - \sin\left[2(p+q)\right]}{2}$$
$$= \frac{\theta_5 + \sin(\theta_6) - \sin(\theta_5 + \theta_6)}{2}.$$

An analogous derivation with $p = \theta_1/2$ and $q = (\theta_5 + \theta_6)/2$ leads to

$$\lambda(S_2) = \frac{\theta_1 + \sin(\theta_5 + \theta_6) - \sin\left[(\theta_1 + \theta_5 + \theta_6)\right]}{2}.$$

Summing up $\lambda(S_1)$ and $\lambda(S_2)$ we obtain the target formula for $\lambda(D_1)$.   □

Proof of Theorem 1 We compute Primus' best reply function. Given t_1, b_1, t_2, b_2, and b, Primus would like to choose t in order to minimise $\lambda(D_1)$. Because of the 1–1 mapping between t and θ_1, we can reformulate this problem as the choice of the optimal angle θ_1 and compute his best reply with respect to θ_1. Differentiating (1) from Lemma 1, we find

$$\frac{\partial \lambda(D_1)}{\partial \theta_1} = \frac{1 - \cos(\theta_1 + \theta_5 + \theta_6)}{2} > 0$$

for any argument, because $0 < |\theta_1 + \theta_5 + \theta_6| < 2\pi$ under the assumption that l and r are interior. Since $\lambda(D_1)$ is (strictly) increasing in θ_1, minimising θ_1 by choosing $t = t_1$ is a dominant strategy for Primus. By a similar argument, $b = b_2$ is a dominant strategy for Secunda. Thus, the unique Nash equilibrium (in dominant strategies) is $(t^*, b^*) = (t_1, b_2)$. $\qquad \square$

Proof of Proposition 2

We use the same notation of the previous proof. Hence, $\tau = \widehat{t_1 o t_2} = \theta_1 + \theta_2$ and $\beta = \widehat{b_1 o b_2} = \theta_4 + \theta_5$. Moreover, $\theta_R = \theta_3$ and $\theta_L = \theta_6$.

Proof The thick line defining the common categorisation divides the disagreement region into two sectors $S_1(t_1 t_2 b_2)$ and $S_2(b_2 b_1 t_1)$. The area $\lambda(S_1)$ is the difference between the areas of the circular segment from t_1 to b_2 with central angle $(\tau + \theta_3)$ and of the circular segment from t_2 to b_2 with central angle θ_3. Hence,

$$\lambda(S_1) = \frac{\tau + \sin \theta_3 - \sin(\tau + \theta_3)}{2}.$$

Similarly,

$$\lambda(S_2) = \frac{\beta + \sin \theta_6 - \sin(\beta + \theta_6)}{2}.$$

Note that $(\tau + \theta_3) + (\beta + \theta_6) = 2\pi$; consequently, $\sin(\tau + \theta_3) = -\sin(\beta + \theta_6)$.

Clearly, Primus is stronger if and only if $\lambda(S_1) - \lambda(S_2) \geq 0$. The sign of the difference

$$\lambda(S_1) - \lambda(S_2) = \frac{\tau - \beta + \sin \theta_3 - \sin \theta_6 - 2\sin(\tau + \theta_3)}{2} \qquad (4)$$

is not trivial. We distinguish two cases and study such sign.

(1) Assume $\tau + \theta_3 \geq \pi \geq \beta + \theta_6$. We consider two sub-cases, depending on the sign of $\theta_6 - \theta_3$. Let us begin with $\theta_6 \geq \theta_3$. We have

$$\begin{aligned}
\lambda(S_1) - \lambda(S_2) &= \frac{\tau - \beta + \sin \theta_3 - \sin \theta_6 - 2\sin(\tau + \theta_3)}{2} \\
&= \frac{2(\tau + \theta_3 - \pi) + [(\theta_6 - \sin \theta_6) - (\theta_3 - \sin \theta_3)] - 2\sin(\tau + \theta_3)}{2}.
\end{aligned}$$

Since $\tau + \theta_3 \geq \pi$ by assumption, the first and the last term in the numerator are positive. Moreover, as the function $x - \sin x$ is increasing on $(0, \pi)$, the term in square brackets is also positive. Hence, $\lambda(S_1) - \lambda(S_2) \geq 0$.

Consider now the sub-case $\theta_6 < \theta_3$. Decomposing S_1 into the circular segment from t_1 to t_2 with central angle τ and the triangle $T(t_1 t_2 b_2)$, we obtain

$$\lambda(S_1) = \frac{\tau - \sin \tau}{2} + 2 \sin\left(\frac{\theta_3}{2}\right) \sin\left(\frac{\tau + \theta_3}{2}\right) \sin\left(\frac{\tau}{2}\right),$$

and similarly,

$$\lambda(S_2) = \frac{\beta - \sin \beta}{2} + 2 \sin\left(\frac{\theta_6}{2}\right) \sin\left(\frac{\beta + \theta_6}{2}\right) \sin\left(\frac{\beta}{2}\right).$$

Hence,

$$\lambda(S_1) - \lambda(S_2) = \frac{(\tau - \sin \tau) - (\beta - \sin \beta)}{2}$$
$$+ 2 \sin\left(\frac{\tau + \theta_3}{2}\right) \left[\sin\left(\frac{\theta_3}{2}\right) \sin\left(\frac{\tau}{2}\right) - \sin\left(\frac{\theta_6}{2}\right) \sin\left(\frac{\beta}{2}\right) \right].$$

The first term is positive by the increasing monotonicity of the function $(x - \sin x)$ on $(0, \pi)$. We claim that the second term is also positive. If $\theta_3 \leq \pi$, this follows because $\sin x$ is increasing in $(0, \pi/2)$, and thus $\sin(\theta_3/2) \sin(\tau/2) \geq \sin(\theta_3/2) \sin(\beta/2) \geq \sin(\theta_6/2) \sin(\beta/2)$. If $\theta_3 > \pi$, then $\theta_6 \leq \tau + \beta + \theta_6 = 2\pi - \theta_3 < \pi$; thus, $\sin(\theta_6/2) \leq \sin(\pi - \theta_3/2) = \sin(\theta_3/2)$, which suffices to establish the claim. From the positivity of the two terms, we conclude that $\lambda(S_1) \geq \lambda(S_2)$.

(2) Assume $\tau + \theta_3 < \beta + \theta_6$. Since by assumption $\tau \geq \beta$, we have $\theta_6 \geq \theta_3$. By (4), using the identity $\tau + \beta + \theta_3 + \theta_6 = 2\pi$, we have

$$2 \left[\lambda(S_1) - \lambda(S_1)\right] = \tau - \beta + \sin\theta_3 + \sin(\tau + \beta + \theta_3) - 2 \sin(\tau + \theta_3)$$

and it suffices to study the sign of the right-hand term. Fix t_2 and b_2. Given τ in $(0, \pi)$, consider the function $f(\beta) = \tau - \beta + \sin\theta_3 + \sin(\tau + \beta + \theta_3) - 2 \sin(\tau + \theta_3)$ for β in $(0, \pi)$. Since $f'(\beta) = -1 + \cos(\tau + \beta + \theta_3) < 0$, the function is strictly decreasing on $[0, \tau]$. Moreover,

$$f(0) = \tau + \sin\theta_3 - \sin(\tau + \theta_3) = [(\tau + \theta_3) - \sin(\tau + \theta_3)] - (\theta_3 - \sin\theta_3) \geq 0$$

by the increasing monotonicity of $(x - \sin x)$ on $[0, \pi]$. Finally, we have

$$f(\tau) = \sin\theta_3 + \sin(\theta_3 + 2\tau) - 2\sin(\tau + \theta_3)$$
$$= \sin(\theta_3) + [\sin(\theta_3)\cos(2\tau) + \cos(\theta_3)\sin(2\tau)]$$
$$- 2[\sin(\tau)\cos(\theta_3) + \cos(\tau)\sin(\theta_3)]$$
$$= \sin(\theta_3)[1 + \cos(2\tau) - 2\cos\tau] + \cos(\theta_3)[\sin(2\tau) - 2\sin\tau]$$

Fig. 8 Visual aids for the proof of Theorem 3
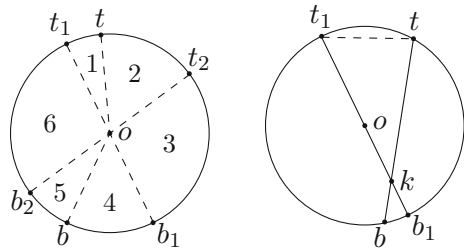
Using the identities $\cos(2\tau) = 2\cos^2\tau - 1$ and $\sin(2\tau) = 2\sin\tau\cos\tau$, we obtain

$$f(\tau) = 2\left[\cos\tau - 1\right]\sin(\tau + \theta_3) \leq 0.$$

By the intermediate value theorem, there exists a unique $\overline{\beta}$ in $[0, \tau]$ such that $f(\overline{\beta}) = 0$. For $\beta \leq \overline{\beta}$, $\lambda(S_1) \geq \lambda(S_2)$ and Primus is stronger. For $\beta > \overline{\beta}$, the opposite inequality holds and Secunda is stronger. □

4.1 Proof of Theorem 3

Similarly to the above (except for switching b_1 and b_2), the endpoints (t_i, b_i) for the two agents' chords and their choices for t and b identify six sectors. Proceeding clockwise, these are numbered from 1 to 6 on the left-hand side of Fig. 8.

For each sector i, we denote its central angle by θ_i. The notation is similar, except that now $\theta_3 = \widehat{t_2ob_1}$, $\theta_4 = \widehat{b_1ob}$, $\theta_5 = \widehat{bob_2}$, and $\theta_6 = \widehat{b_2ot_1}$. Recall that $\tau = \theta_1 + \theta_2$ and $\beta = \theta_4 + \theta_5$; moreover, since the categorisations are characterised by diameters, $\tau = \beta$. The following lemma characterises the disagreement area of each player as a function of the six central angles.

Lemma 3 *The disagreement areas for Primus and Secunda are, respectively:*

$$\lambda(D_1) = \frac{\theta_1 - \sin\theta_1}{2} + \frac{\theta_4 - \sin\theta_4}{2} + 2\cos\left(\frac{\theta_1}{2}\right)\cos\left(\frac{\theta_4}{2}\right)\frac{\sin^2(\theta_1/2) + \sin^2(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)},$$

(5)

and

$$\lambda(D_2) = \frac{\theta_2 - \sin\theta_2}{2} + \frac{\theta_5 - \sin\theta_5}{2} + 2\cos\left(\frac{\theta_2}{2}\right)\cos\left(\frac{\theta_5}{2}\right)\frac{\sin^2(\theta_2/2) + \sin^2(\theta_5/2)}{\sin(\theta_2/2 + \theta_5/2)}.$$

Proof The disagreement region D_1 for Primus can be decomposed into the two sector-like regions $S_1(t_1tk)$ and $S_2(kb_1b)$ as shown on the right-hand side of Fig. 8. We compute the areas $\lambda(S_1)$ and $\lambda(S_2)$, and then add them up to obtain $\lambda(D_1)$.

The region $S_1(t_1tk)$ can be decomposed into two parts: the circular segment from t_1 to t with central angle θ_1, and the triangle $T(t_1tk)$. The area of the circular segment is $(\theta_1 - \sin\theta_1)/2$. The computation of the area of the triangle needs to take into account that the position of k depends on t. We use the ASA formula: given the

length a of one side and the size of its two adjacent angles α and γ, the area is $(a^2 \sin \alpha \sin \gamma)/(2 \sin(\alpha + \gamma))$. We pick $a = \overline{tt_1}$, $\alpha = \widehat{kt_1t}$, and $\gamma = \widehat{t_1tk}$. By the inscribed angle theorem, $\alpha = (\pi - \theta_1)/2$ and $\gamma = (\pi - \theta_4)/2$. Recall that $\overline{tt_1} = 2 \sin(\theta_1/2)$; moreover, $\sin \alpha = \sin((\pi - \theta_1)/2) = \cos(\theta_1/2)$ and, similarly, $\sin \gamma = \cos(\theta_4/2)$. Hence,

$$\lambda(T) = \frac{2\,(\sin(\theta_1/2))^2 \cdot \cos(\theta_1/2) \cdot \cos(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)}.$$

Adding up the two areas, we obtain

$$\lambda(S_1) = \frac{\theta_1 - \sin \theta_1}{2} + \frac{2\,(\sin(\theta_1/2))^2 \cdot \cos(\theta_1/2) \cdot \cos(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)}.$$

By a similar argument,

$$\lambda(S_2) = \frac{\theta_4 - \sin \theta_4}{2} + \frac{2\,(\sin(\theta_4/2))^2 \cdot \cos(\theta_1/2) \cdot \cos(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)}.$$

Summing up $\lambda(S_1)$ and $\lambda(S_2)$ provides the formula for $\lambda(D_1)$. The derivation of a specular formula for $\lambda(D_2)$ is analogous.   □

A direct study of the sign of the derivative $\partial\lambda(D_1)/\partial\theta_1$ is quite involved, but the following lemma greatly simplifies it. An analogous result holds for Secunda.

Lemma 4 *Let* $a = \cos(\theta_4/2)$, $b = \sin(\theta_4/2)$, $c = ab = \sin(\theta_4)/2$, *and* $x = \tan(\theta_1/4)$. *Then*

$$\operatorname{sgn}\left[\frac{\partial\lambda(D_1)}{\partial\theta_1}\right] = \operatorname{sgn}[P(x)], \tag{6}$$

where

$$P(x) = -\left[c\left(1 + x^2\right)^2 - 2\left(\sqrt{2} + 1\right) x \left(1 - x^2\right)\right].$$

Proof Differentiating (5) from Lemma 3 and using a few trigonometric identities, we obtain

$$\begin{aligned}
\frac{\partial\lambda(D_1)}{\partial\theta_1} =& \frac{1 - \cos\theta_1}{2} + \frac{2\sin(\theta_1/2)\cos^2(\theta_1/2)\cos(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)} \\
& - \frac{\cos(\theta_4/2)\left[\sin^2(\theta_1/2) + \sin^2(\theta_4/2)\right]}{\sin^2(\theta_1/2 + \theta_4/2)} \\
& \times \left[\sin(\theta_1)\sin(\theta_1/2 + \theta_4/2) + \cos(\theta_1)\cos(\theta_1/2 + \theta_4/2)\right]
\end{aligned}$$

$$= \sin^2(\theta_1/2) + \frac{\sin(\theta_1)\cos(\theta_1/2)\cos(\theta_4/2)}{\sin(\theta_1/2 + \theta_4/2)}$$
$$- \frac{\cos^2(\theta_4/2)\left[\sin^2(\theta_1/2) + \sin^2(\theta_4/2)\right]}{\sin^2(\theta_1/2 + \theta_4/2)}$$

Let $a = \cos(\theta_4/2)$, $b = \sin(\theta_4/2)$, and $x = \tan(\theta_1/4)$. Recall the double angle formulas $\sin(\theta_1/2) = 2x/(1 + x^2)$ and $\cos(\theta_1/2) = (1 - x^2)/(1 + x^2)$. Then

$$\sin\left(\frac{\theta_1 + \theta_4}{2}\right) = a\left(\frac{2x}{1 + x^2}\right) + b\left(\frac{1 - x^2}{1 + x^2}\right) = \frac{2ax + b(1 - x^2)}{1 + x^2}.$$

Substituting with respect to the new variable x, we find

$$\frac{\partial\lambda(D_1)}{\partial\theta_1} = \left(\frac{2x}{1 + x^2}\right)^2 + \frac{4ax(1 - x^2)^2}{(1 + x^2)^2\left[2ax + b(1 - x)^2\right]} - \frac{a^2\left[4x^2 + b^2(1 + x^2)^2\right]}{\left[2ax + b(1 - x)^2\right]^2}$$
$$= -\frac{N(x)}{(1 + x^2)^2\left[2ax + b(1 - x^2)\right]^2}, \tag{7}$$

where using the identity $a^2 + b^2 = 1$, the polynomial in the numerator can be written as

$$N(x) = a^2\left(1 + x^2\right)^2\left[4x^2 + b^2\left(1 + x^2\right)^2\right] - 4ax\left(1 - x^2\right)^2\left[2ax + b(1 - x^2)\right]$$
$$- 4x^2\left[2ax + b(1 - x^2)\right]^2.$$

Let $c = ab = \sin(\theta_4)/2$ and rewrite $N(x)$ after collecting terms with respect to c:

$$N(x) = c^2\left(1 + x^2\right)^4 - 4cx\left(1 - x^2\right)\left(1 + x^2\right)^2 - 4x^2\left(1 - x^2\right)$$
$$= \left[c\left(1 + x^2\right)^2 - 2x\left(1 - x^2\right)\right]^2 - \left[2\sqrt{2}x\left(1 - x^2\right)\right]^2$$
$$= \left[c\left(1 + x^2\right)^2 - 2\left(\sqrt{2} + 1\right)x\left(1 - x^2\right)\right]$$
$$\times \left[c\left(1 + x^2\right)^2 + 2\left(\sqrt{2} - 1\right)x\left(1 - x^2\right)\right].$$

As both θ_1 and θ_4 are in the open interval $(0, \pi)$ by construction, we have $x = \tan(\theta_1/4) > 0$ and $c = \sin(\theta_4)/2 > 0$; hence, the second term in the multiplication is strictly positive. Returning to (7), this implies

$$\mathrm{sgn}\left[\frac{\partial\lambda(D_1)}{\partial\theta_1}\right] = -\mathrm{sgn}\left[N(x)\right] = \mathrm{sgn}\left[P(x)\right],$$

with $P(x) = -\left[c(1 + x^2)^2 - 2(\sqrt{2} + 1)x(1 - x^2)\right]$, as it was to be shown. $\square$

It is convenient to work with the central angles subtended by the points on the circumference. Recall that, given t_1, t_2, b_1, and b_2, Primus and Secunda simultaneously choose t and b, respectively. Then Secunda's choice of b is in a 1-1 mapping with the angle $\theta_5 = \widehat{b_2 o b}$, while Primus' choice of t has a similar relation to $\theta_1 = \widehat{t_1 o t}$.

The following lemma characterizes Primus' best reply using the central angles θ_1 and θ_5, rather than the endpoints t and b. As it turns out, such best reply is always unique; hence, with obvious notation, we denote it as the function $\theta_1 = r_1(\theta_5)$. Correspondingly, let $\theta_5 = r_2(\theta_1)$ be the best reply function for Secunda. Finally, recall our assumption that the individual categorisations are supported by diameters: this implies that the two angular distances $\tau = \theta_1 + \theta_2$ and $\beta = \theta_4 + \theta_5$ are equal with $0 \le \tau = \beta < \pi$; moreover, players' initial positions have the same strength and the game is symmetric.

Lemma 5 *The best reply functions for the two players are*

$$r_1(\theta_5) = \arcsin\left(\frac{\sin(\beta - \theta_5)}{\sqrt{2} + 1}\right) \quad and \quad r_2(\theta_1) = \arcsin\left(\frac{\sin(\tau - \theta_1)}{\sqrt{2} + 1}\right),$$

with $0 \le \theta_5 \le \beta$ and $0 \le \theta_1 \le \tau$.

Proof Consider Primus. (The argument for Secunda is identical.) For any θ_5 in $[0, \beta]$, we search which value of θ_1 in $[0, \tau]$ minimises $\lambda(D_1)$. We distinguish two cases.

First, suppose $\theta_5 = \beta$. Then $\theta_4 = 0$ and $\lambda(D_1) = (\theta_1 + \sin \theta_1)/2$. As this function is increasing in θ_1, the optimal value is $\theta_1^* = 0$.

Second, suppose $\theta_5 < \beta$. We begin by finding the stationary points of $\lambda(D_1)$. Recall that we let $x = \tan(\theta_1/4)$. By Lemma 4, $\partial \lambda(D_1)/\partial \theta_1 = 0$ if and only if $P(x) = 0$; that is, if and only if

$$c = \frac{2\left(\sqrt{2} + 1\right) x \left(1 - x^2\right)}{\left(1 + x^2\right)^2}.$$

Replacing the double angle formulæ $\sin(\theta_1/2) = 2x/(1 + x^2)$ and $\cos(\theta_1/2) = (1 - x^2)/(1 + x^2)$, we obtain

$$c = \left(\sqrt{2} + 1\right) \sin\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_1}{2}\right) = \left(\sqrt{2} + 1\right) \frac{\sin \theta_1}{2}.$$

On the other hand, since $c = (\sin \theta_4)/2$ by definition and $\theta_4 + \theta_5 = \beta$, this yields

$$\sin \theta_1 = \frac{\sin \theta_4}{\sqrt{2} + 1} = \frac{\sin(\beta - \theta_5)}{\sqrt{2} + 1}.$$

Since $\theta_5 \in [0, \beta]$, the only solutions to this equation are the supplementary angles θ_1' and $\theta_1'' = \pi - \theta_1'$ with

$$\theta_1' = \arcsin\left(\frac{\sin(\beta - \theta_5)}{\sqrt{2} + 1}\right) < \frac{\pi}{2} < \pi - \theta_1' = \theta_1''.$$

These are the stationary points for $\lambda(D_1)$.

Clearly, $\theta_1' \geq 0$. We claim that $\theta_1' < \tau$. If $\pi/2 \leq \tau$, this is obvious. Suppose instead $\tau < \pi/2$. Since $\theta_4 < \beta = \tau < \pi/2$, we have $\sin\theta_1' = (\sqrt{2} - 1)\sin(\theta_4) < \sin\theta_4 < \sin\tau$ and thus $\theta_1' < \tau$. We conclude that the stationary point θ_1' belongs to the interval $[0, \tau]$.

For $\theta_1 = 0$, we have $x = 0$ and $P(x)|_{x=0} = -c = -(\sin\theta_4)/2 < 0$. Therefore, we have by continuity that $P(x)$ changes sign from negative to positive in θ_1' and from positive to negative in θ_1''. By Lemma 4, this implies that the only local minimisers for $\lambda(D_1)$ in the compact interval $[0, \tau]$ are $\theta = \theta_1'$ and $\theta = \tau$. Comparing the corresponding values for $\lambda(D_1)$, we find

$$\lambda(D_1)|_{\theta_1=\theta'} < \lambda(D_1)|_{\theta_1=0} < \lambda(D_1)|_{\theta_1=\tau},$$

where the first inequality follows from the (strict) negativity of $\partial\lambda(D_1)/\partial\theta_1$ in $[0, \theta']$ and the second inequality from a direct comparison. Hence, the global minimiser is θ'. Combining the two cases, it follows that, for any θ_5 in $[0, \beta]$, the unique best reply is $r_1(\theta_5) = \arcsin\left[\sin\left(\beta - \theta_5\right)/(\sqrt{2} + 1)\right]$. $\qquad\square$

Proof of Theorem 3 A Nash equilibrium is any fixed point (θ_1, θ_5) of the map

$$\begin{pmatrix} \theta_1 \\ \theta_5 \end{pmatrix} = \begin{pmatrix} r_1(\theta_5) \\ r_2(\theta_1) \end{pmatrix}$$

from $[0, \tau] \times [0, \beta]$ into itself. Substituting from Lemma 5 and using $\tau = \beta$, we obtain the system of equations

$$\begin{cases} \sin(\theta_1) = \dfrac{\sin(\tau - \theta_5)}{\sqrt{2} + 1} \\[3mm] \sin(\theta_5) = \dfrac{\sin(\tau - \theta_1)}{\sqrt{2} + 1} \end{cases} \qquad (8)$$

Multiplying across gives

$$\sin(\theta_1)\sin(\tau - \theta_1) = \sin(\theta_5)\sin(\tau - \theta_5);$$

or, using the prosthaphaeresis formula,

$$\cos(2\theta_1 - \tau) - \cos\tau = \cos(2\theta_5 - \tau) - \cos\tau$$

from which we get that the only two possible solutions in $[0, \tau]$ are

$$\theta_1 = \theta_5 \quad \text{or} \quad \theta_1 = \tau - \theta_5.$$

When $\tau > 0$, the second possibility can be discarded because, when replaced in (8), it would yield the contradiction $\theta_1 = \theta_5 = \tau = 0$. (When $\tau = 0$, we trivially obtain $\theta_1 = -\theta_5 = \tau = 0$ as in the first case.) Hence, we are left with $\theta_1 = \theta_5$.

Substituting in the first Eq. of (8), we obtain

$$\sin(\theta_1) = \frac{\sin(\tau - \theta_1)}{\sqrt{2} + 1} = \frac{\sin \tau \cos \theta_1 - \cos \tau \sin \theta_1}{\sqrt{2} + 1}.$$

As $0 \leq \theta_1 < \pi/2$, dividing by $\cos \theta_1$ yields

$$\tan(\theta_1) = \frac{\sin \tau}{\sqrt{2} + 1 + \cos \tau}$$

and the result follows.   □

4.2 Proof of Proposition 4

Proof Recall that the payoff for an agent is the opposite of the area of the disagreement region. Consider Primus. (The proof for Secunda is analogous.) Let D^* and D^s be the region of disagreement between Primus' and the common categorisation at the equilibrium and, respectively, at the Nash cooperative solution. For $\tau = 0$, $D^* = D^s$. Hence, we assume $\tau \neq 0$ and show that $\lambda(D^*) - \lambda(D^s) > 0$.

At the Nash bargaining solution, $\theta_1^s = \theta_2^s = \tau/2$; replacing these into (6), we find $\lambda(D^s) = \tau/2$. At the Nash equilibrium, $\theta_1^* = \theta_5^*$ and thus $\theta_4^* = \tau - \theta_1^*$; substituting these into (6) and dropping superscripts and subscripts for simplicity, we obtain

$$\lambda(D^*) = \frac{\tau}{2} - \left[\frac{\sin \theta + \sin(\tau - \theta)}{2}\right]$$
$$+ 2\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{\tau - \theta}{2}\right)\frac{\sin^2(\theta/2) + \sin^2((\tau - \theta)/2)}{\sin(\tau/2)}.$$

Hence, using standard trigonometric identities,

$$\lambda(D^*) - \lambda(D^s) = -\left[\frac{\sin \theta + \sin(\tau - \theta)}{2}\right]$$
$$+ 2\cos\left(\frac{\theta}{2}\right)\cos\left(\frac{\tau - \theta}{2}\right)\frac{\sin^2(\theta/2) + \sin^2((\tau - \theta)/2)}{\sin(\tau/2)}$$
$$= -\sin\left(\frac{\tau}{2}\right)\cos\left(\theta - \frac{\tau}{2}\right)$$
$$+ \left[\frac{1}{\sin(\tau/2)}\right]\left[\cos\left(\frac{\tau}{2}\right) + \cos\left(\theta - \frac{\tau}{2}\right)\right]\left[1 - \frac{\cos \theta}{2} - \frac{\cos(\tau - \theta)}{2}\right]$$
$$= -\left[\frac{1}{\sin(\tau/2)}\right]\left[1 - \cos^2\left(\frac{\tau}{2}\right)\right]\cos\left(\theta - \frac{\tau}{2}\right)$$
$$+ \left[\frac{1}{\sin(\tau/2)}\right]\left[\cos\left(\frac{\tau}{2}\right) + \cos\left(\theta - \frac{\tau}{2}\right)\right]$$

$$\times \left[1 - \cos\left(\frac{\tau}{2}\right) \cos\left(\theta - \frac{\tau}{2}\right) \right]$$
$$= \left[\frac{\cos(\tau/2)}{\sin(\tau/2)} \right] \sin^2\left(\theta - \frac{\tau}{2}\right),$$

from which we obtain

$$\mathrm{sgn}\left[\lambda(D^*) - \lambda(D^s) \right] = \mathrm{sgn}\left[\tan(\tau/2) \right].$$

Since $0 < \tau < \pi$, $\tan(\tau/2) > 0$, and the claim follows.   □

References

Benz, A., Jager, G., & van Rooy, R. (Eds.). (2005). *Game theory and pragmatics*. Basingstoke: Palgrave Macmillan.

Clark, R. (2012). *Meaningful games: Exploring language with game theory*. Cambridge: The MIT Press.

Cohen, H., & Lefebvre, C. (Eds.). (2005). *Handbook of categorization in cognitive science*. Amsterdam: Elsevier.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge: The MIT Press.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge: The MIT Press.

Gärdenfors, P. (2014a). Levels of communication and lexical semantics. Synthese, accepted for publication.

Honkela, T., Könönen, V., Lindh-Knuutila, T., & Paukkeri, M.-S. (2008). Simulating processes of concept formation and communication. *Journal of Economic Methodology*, *15*, 245–259.

Jäger, G. (2007). The evolution of convex categories. *Linguistics and Philosophy*, *30*, 551–564.

Jäger, G., & Van Rooij, R. (2007). Language structure: Psychological and social constraints. *Synthese*, *159*, 99–130.

Jäger, G., Metzger, L. P., & Riedel, F. (2011). Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and Economic Behavior*, *73*, 517–537.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, *4*, 71–84.

Kruschke, J. K. (2008). Models of categorization. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (pp. 267–301). New York: Cambridge University Press.

Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. Oxford: Oxford University Press.

Mervis, C., & Rotsch, E. (1981). Categorisation of natural objects. *Annual Review of Psychology*, *32*, 89–115.

Parikh, P. (2010). *Language and equilibrium*. Cambridge: The MIT Press.

Rotsch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology*, *104*, 192–233.

Warglien, M., & Gärdenfors, P. (2013). Semantics, conceptual spaces, and the meeting of minds. *Synthese*, *190*, 2165–2193.

Warglien, M., & Gärdenfors, P. (2015). Meaning negotiation. In F. Zenker & P. Gärdenfors (Eds.), *Applications of conceptual spaces: The case for geometric knowledge representation* (pp. 79–94). Dordrecht: Springer.

Wernerfelt, B. (2004). Organizational languages. *Journal of Economics and Management Strategy*, *13*, 461–472.